Recherche et évaluation en comportement

Ministère des Communications

RECHERCHE EN COMPORTEMENT SUR TÉLIDON 2

LA CONCEPTION D'INDEX RAMIFIÉS POUR LE VIDÉOTEX

Mai 1981

Les copies de

RECHERCHE EN COMPORTEMENT SUR TÉLIDON 2
LA CONCEPTION D'INDEX RAMIFIÉS POUR LE VIDÉOTEX

sont disponibles gratuitement à la:

Direction de l'information
Ministère des Communications
300 rue Slater
Ottawa, Canada
K1A 0C8

Numéro du rapport DOC-TBR-DBRE-81-2-F

RECHERCHE EN COMPORTEMENT SUR TÉLIDON 2
MINISTÈRE DES COMMUNICATIONS
OTTAWA
MAI 1981

LA CONCEPTION D'INDEX RAMIFIÉS POUR LE VIDÉOTEX

Préface

L'élaboration de bases de données vidéotex fait appel à des structures hiérarchisées ou arborescentes. Les rapports présentés dans le présent document indiquent clairement qu'un index mal conçu peut rendre la recherche documentaire difficile, voire impossible pour certains utilisateurs, même quand la base de données contient les informations recherchées.

Les recherches consacrées à cette méthode d'extraction sont peu nombreuses et on n'a pas encore élaboré les lignes directrices dont les concepteurs d'index pourraient s'inspirer pour la construction d'index faciles à consulter.

Les présents rapports émanent des laboratoires du Groupe de recherche en comportement du ministère des Communications et se veulent une première étape dans l'élaboration d'un ensemble de lignes directrices s'appliquant à la construction d'index ramifiés. Ils indiquent que pendant une recherche documentaire portant sur un point précis, l'utilisateur commet effectivement beaucoup d'erreurs dans le choix des rubriques. Un certain nombre de recommandations visant à améliorer la structure arborescente en découlent. En outre, les méthodes employées pour l'essai des index ramifiés peuvent être adaptées à l'essai d'index mis au point dans l'entreprise privée. Le ministère des Communications poursuit ses recherches sur la question. Le présent document contient quatre rapports:

1. L'efficacité d'un index ramifié quand l'existence de données est incertaine

   Ce rapport présente une étude du processus de recherche de participants incertains de trouver dans la base de données la réponse à leurs questions, situation semblable à celle dans laquelle se trouve un utilisateur normal. Les résultats de l'expérience menée à l'aide d'un index simulé de la base de données du Télidon permettent de constater non seulement que l'on commet des erreurs dans le choix des rubriques, mais aussi que la plupart de celles-ci portent sur les pages des deux premiers niveaux de la structure arborescente. Ils montrent en outre que les utilisateurs mettent fin à leurs recherches avant de parvenir aux données en mémoire.

2. L'utilisation d'un index ramifié contenant trois défauts de conception.

Ce rapport présente une expérience dans laquelle des défauts de conception ont été ajoutés à dessein à la base de données. On a encore une fois utilisé un index simulé de la base de données de démonstration du Télidon. Les défauts étaient les suivants: des données placées dans la mauvaise catégorie, deux rubriques synonymes dans la même page et des titres de rubriques imprécis. Les participants ont jugé que le défaut le plus sérieux prolongeant le plus le temps de recherche était l'insertion des données sous la mauvaise rubrique. Les autres défauts ont aussi causé une baisse du rendement.

3. Étude de l'efficacité d'utilisation d'un système de recherche documentaire Télidon.

On trouve dans ce rapport les résultats d'une expérience faisant appel pour la première fois à une base de données Télidon réelle (il s'agit à nouveau de la base de démonstration du Télidon du ministère des Communications). Beaucoup d'erreurs ont été commises au cours de la recherche documentaire. En fait, le taux d'erreurs était supérieur à celui relevé dans les expériences antérieures au cours desquelles on utilisait une base de données simulée. Comme dans le cas des expériences antérieures, beaucoup d'erreurs portaient sur les deux premiers niveaux de l'arbre.

4. La conception d'index ramifiés du vidéotex: utilisation de descripteurs et amélioration des index page par page.

Ce rapport expose les résultats de deux expériences conçues pour évaluer les effets de l'addition de descripteurs (brèves descriptions) aux rubriques, mettre à l'épreuve la page de base ou la page un de l'index de différentes façons et démontrer des méthodes d'essai et de modification d'index ramifiés.

On remarquera que les expériences décrites dans ce rapport ont été menées à l'aide de l'index de la base de données de démonstration du Télidon du ministère des Communications qui en était encore à un stade très peu avancé de son développement. La base de données a subi d'importantes modifications depuis ce temps. On ne doit donc pas s'appuyer sur les chiffres absolus mentionnés, par exemple les taux d'erreurs, pour étendre les conclusions obtenues à d'autres index du Télidon. Les expériences permettent toutefois aux auteurs de tirer des conclusions sur les principes de la conception d'un index ramifié. Un résumé de ces conclusions est intégré aux lignes directrices exposées ci-dessous.

Lignes directrices ayant trait à la construction d'index ramifiés: résumé des conclusions découlant des expériences décrites dans les rapports mentionnés ci-dessus.

1. L'essai d'un index ramifié par des non-initiés (personnes ne connaissant pas la base de données) peut améliorer celui-ci de façon marquée. Les méthodes qui ont servi aux expériences décrites dans les présents rapports pourraient être adaptées à l'intention des fournisseurs commerciaux de base de données.

   - Lorsqu'on dispose de ressources restreintes, la méthode la plus efficace consiste à faire l'essai des deux premiers niveaux de la structure arborescente.

   - Les essais devraient être menés à l'aide de la population cible.

   - Les essais peuvent être effectués à l'aide de la méthode page par page (c'est-à-dire concevoir plusieurs versions d'une même page et mesurer le rendement et les préférences).

   - D'après les données dont nous disposons actuellement, on doit mesurer tant les préférences que le rendement.

2. On peut améliorer les pages d'index en éliminant les possibilités d'erreurs et en tenant compte des préférences des utilisateurs.

   - Le fait d'insérer les données sous une mauvaise rubrique constitue un défaut de conception très sérieux qui porte l'utilisateur à commettre des erreurs ou le rend absolument incapable de trouver les renseignements recherchés.

   - Des titres de rubriques ambigus ou imprécis ou des titres de rubriques synonymes sont une cause d'erreurs.

   - Les utilisateurs préfèrent des quantités égales d'informations sous chaque rubrique (et non 80 pour cent sous une seule rubrique).

   - Les utilisateurs préfèrent retrouver plus de quatre rubriques par page, ce qui leur permet probablement de trouver les renseignements recherchés plus rapidement que lorsque les pages n'en contiennent que deux ou trois.

- Les utilisateurs préfèrent un classement rationnel des rubriques, c'est-à-dire par ordre alphabétique, ou selon la fréquence de consultation des rubriques.

Dorothy Phillips
Recherche et évaluation en comportement
Ministère des Communications
300, rue Slater
Ottawa (Ontario) K1A 0C8
Mai 1981

RECHERCHE EN COMPORTEMENT SUR TÉLIDON 2

MAI 1981

LA CONCEPTION D'INDEX RAMIFIÉS DU VIDÉOTEX

Table des matières — Page

Préface - Dorothy Phillips

I L'efficacité d'un index ramifié quand l'existence de données est incertaine - Thomas Whalen et Susane Latrémouille..... 1

II L'utilisation d'un index ramifié contenant trois défauts de conception - Thomas Whalen et Candy Mason..... 17

III Étude de l'efficacité d'utilisation d'un système de recherche documentaire Télidon - Scott A. McEwen..... 39

IV La conception d'index ramifiés du vidéotex: utilisation de descripteurs et amélioration des index page par page - Susane Latrémouille et Eric Lee..... 75

CHAPITRE 1

L'EFFICACITÉ D'UN INDEX RAMIFIÉ LORSQUE L'EXISTENCE DE DONNÉES EST INCERTAINE

Thomas Whalen et Susane Latrémouille

L'EFFICACITÉ D'UN INDEX RAMIFIÉ LORSQUE L'EXISTENCE DE DONNÉES EST INCERTAINE

RÉSUMÉ

On a demandé à huit personnes de trouver 16 renseignements à l'aide d'un système à structure arborescente mis en oeuvre sur un ordinateur en temps partagé. On a informé les participants que certains des renseignements n'étaient pas disponibles. Dans l'ensemble, la probabilité que le sujet fasse un choix inexact dans une page d'index était de 0,15. La plupart des erreurs commises portaient sur les deux premiers niveaux de l'arbre. Ces résultats confirment ceux d'études antérieures. Cette expérience a aussi démontré que la plupart des recherches se terminent par l'obtention du renseignement recherché après une consultation de 4,52 pages d'index en moyenne. S'ils ne trouvaient pas le renseignement voulu après avoir consulté 8,78 pages en moyenne, les participants mettaient fin à la recherche. Nous avons donc conclu que les gens cesseront très probablement leurs recherches avant d'avoir trouvé un renseignement existant dans la base de données plutôt que d'effectuer des recherches étendues. Ceci laisse donc entendre qu'un utilisateur interrompra sa recherche dans la base de données du Télidon si la structure arborescente comporte des défauts de conception.

INTRODUCTION

L'extraction de données d'une base de données générale par des personnes expérimentées est une des premières applications du Télidon. Même s'il est généralement convenu que la construction d'une structure arborescente est un moyen d'atteindre ce but, on dispose de peu de données montrant la facilité d'utilisation d'un système de ce genre, particulièrement lorsque l'utilisateur ignore si les renseignements qu'il souhaite obtenir sont dans la base de données.

La première partie du présent exposé décrit les résultats d'une expérience portant sur l'utilisation d'un système à structure arborescente lorsque certains des renseignements que cherche l'utilisateur n'existent pas. La seconde partie présente une comparaison entre ces résultats et ceux d'une expérience antérieure. Une description détaillée de la méthode utilisée aux fins de l'expérience est incluse en annexe.

PROBABILITÉ D'UNE RECHERCHE FRUCTUEUSE

On a demandé à chacun des participants d'accomplir 16 tâches dont chacune consistait à trouver un renseignement précis en faisant des choix dans les pages d'un index ramifié affichées de façon séquentielle sur un terminal d'ordinateur. Vingt-cinq des problèmes étaient insolubles parce que les renseignements recherchés n'étaient pas dans la base de données. Les participants ont résolu 86,5% des problèmes solubles. La probabilité que les participants trouvent les renseignements recherchés est donc très élevée, même quand ceux-ci savent qu'une recherche prolongée peut se révéler infructueuse.

Il est étonnant de constater que ces personnes ont si rarement mis fin à une recherche avant de trouver le renseignement recherché. On peut toutefois attribuer ceci au fait que les participants ont réussi à trouver le renseignement rapidement quand il était disponible. Dans le cas des problèmes solubles, une recherche parfaite exigeait la consultation de 3,58 pages en moyenne. Les recherches fructueuses ont exigé la consultation de 4,52 pages en moyenne. Les participants ont donc trouvé le renseignement recherché presque aussi rapidement que possible. En moyenne, ils n'ont consulté qu'une page supplémentaire.

Pour le premier problème donnant lieu à une recherche infructueuse, les participants ont consulté en moyenne 21,87 pages. Pour la deuxième recherche se terminant par un échec, ce chiffre tombe à 10,12. Ceci laisse entendre qu'un processus d'apprentissage était en cours. Afin de tirer des conclusions sur la stratégie normale de recherche d'un participant, nous avons exclu de l'analyse exposée ci-dessous le premier problème non résolu. Ce problème exclu, les participants ont donc consulté en moyenne 8,78 pages avant de mettre fin à une recherche infructueuse. Ils ont donc consulté deux fois plus de pages lorsqu'ils ne pouvaient trouver le renseignement recherché que lorsqu'ils réussissaient à le trouver.

Dans l'ensemble, le système à index ramifié est donc pratique quand l'existence de renseignements est incertaine parce que les utilisateurs trouvent le renseignement rapidement et non parce qu'ils sont prêts à mener des recherches exhaustives dans la base de données. On peut donc affirmer que la conception de la base de données est importante. Si les utilisateurs ne pouvaient trouver les renseignements recherchés à l'aide d'un système à structure arborescente mal conçu, ils concluraient rapidement que les renseignements ne sont pas disponibles.

PAGES A TAUX D'ERREURS ÉLEVÉ

Dans tous les cas où un participant pouvait faire un bon choix, il l'a fait 85,3% du temps. Cette probabilité de choix exact semble élevée, mais on doit en tenir compte dans le contexte d'une recherche dans un système à structure arborescente. Pour trouver un renseignement dans la base de données, on doit consulter un certain nombre de pages. Au fur et à mesure de l'élargissement de la base de données, le nombre de pages à consulter s'accroît. Si la probabilité d'erreur est de 15% par page, la probabilité de faire dix bons choix consécutifs est inférieure à 20%.

Cette analyse suppose que les erreurs sont réparties de façon uniforme entre toutes les pages. La plupart des erreurs se produisent cependant dans un petit nombre de pages. Dans la présente expérience, les erreurs pouvaient porter sur 25 pages, mais n'ont porté que sur 11 pages. La plupart de ces pages ont fait l'objet d'une ou deux erreurs seulement, de sorte que, dans l'ensemble, 45 des 57 erreurs se sont produites dans seulement 16% des pages. Une de ces pages était celle qui est située à la racine de l'arbre et les trois autres, celles qui la suivent immédiatement. Ces résultats nous portent à croire que le fait d'accroître la profondeur de l'arbre en ajoutant des pages d'index dans sa partie inférieure pourrait ne pas augmenter la difficulté de la recherche documentaire. D'autres recherches expérimentales devront être effectuées en vue de vérifier cette possibilité.

COMPARAISON AVEC DES RECHERCHES ANTÉRIEURES

Dans un mémoire technique rédigé antérieurement (1980), Lee et Latrémouille décrivent une recherche dans des pages d'index disposées selon une structure arborescente. L'étude a été menée à l'aide de pages de menu dactylographiées sur des fiches de 5 pouces sur 7 pouces. On a demandé à chaque participant de trouver des renseignements à l'aide de cet index en vue d'évaluer la structure arborescente. Les deux chercheurs ont aussi recueilli les commentaires des participants afin de déterminer les causes des erreurs commises.

La façon d'effectuer les recherches diffère, dans cette méthode expérimentale, de celle qui a été employée dans l'expérience utilisant le Télidon. D'abord les participants savaient d'avance que les renseignements recherchés existaient réellement dans la base de données. Cependant, lorsqu'on effectue une recherche dans la base de données du Télidon, on ignore le plus souvent, en commençant, si la base de données contient ou non les renseignements désirés. Ainsi, après avoir vu le résultat de chacun de ses choix, l'utilisateur doit aussi prendre une décision concernant l'opportunité de continuer la recherche.

Deuxièmement, les personnes participant à l'expérience menée par Lee et Latrémouille ont été informées immédiatement de l'exactitude ou de l'inexactitude de leur choix. Grâce à ce renseignement, les participants disposaient des données dont ils avaient besoin pour résoudre le problème qui leur était posé. Cette façon de procéder permettait aux participants de ne pas commettre la même erreur deux fois de suite même s'ils pouvaient commettre des erreurs différentes de la même nature à plusieurs reprises. Un utilisateur du Télidon fera toutefois toujours le même choix erroné dans une page qu'il ne comprend pas.

Lee et Latrémouille ont donc mené leur expérience dans des circonstances différentes de celles qui existent dans le cours normal de l'utilisation du Télidon. Pour mieux représenter la situation, la présente étude a simulé la base de données de démonstration du Télidon sur un gros ordinateur en temps partagé. Les pages d'index étaient visualisées sur un terminal vidéo ordinaire et les participants choisissaient les rubriques à l'aide d'un clavier placé sur le terminal. Les pages d'index utilisées au cours de l'expérience de Lee et Latrémouille étaient essentiellement les mêmes que celles qui ont été utilisées au cours de la présente expérience, mais les tâches à accomplir au cours de la première expérience n'avaient aucun rapport avec celles qui ont été choisies pour la présente expérience.

Étant donné l'ampleur des différences entre les deux expériences, il est étonnant de constater la mesure de ressemblance entre les résultats. Lee et Latrémouille ont signalé une probabilité d'erreur de 14% alors que la présente expérience a fait état d'une probabilité d'erreur de 14,7%. La similitude de ces résultats est sans aucun doute accidentelle, mais elle indique néanmoins que la différence entre les méthodes expérimentales n'a pas d'effet marqué sur les résultats. Cette constatation importe en ce sens qu'elle permet de valider la méthode utilisée par Lee et Latrémouille: il est effectivement possible de faire l'essai de pages d'index en les dactylographiant sur des fiches et en demandant à des personnes d'y choisir des rubriques. Il est inutile de se donner la peine d'incorporer les pages d'index dans une base de données complète.

Ces deux expériences montrent également que les participants ont fait preuve d'une grande cohérence, la majorité des erreurs ayant été commises aux deux premiers niveaux de l'arbre. Dans l'ensemble, Lee et Latrémouille ont découvert que 80% des erreurs portaient sur six des 79 pages de menu et 58% sur les deux premiers niveaux de l'arbre.

La même conclusion générale se dégage donc des deux études: un système à index hiérarchisé ne sera utile que dans la mesure où il est bien conçu. Au fur et à mesure de l'élargissement de la base de données, il devient plus important de diminuer le plus possible la probabilité que l'utilisateur commette une erreur dans chacune des pages.

DOCUMENTATION

Lee E., et Latrémouille, S., Évaluation de la structure arborescente de l'information dans la base de données du Télidon. Recherche en comportement sur Télidon 1. Ministère des Communications, février 1980.

ANNEXE: LA MÉTHODE EXPÉRIMENTALE

SUJETS

Aucun des huit volontaires qui ont participé à la présente étude n'avait utilisé le Télidon antérieurement.

MÉTHODE

On a expliqué aux participants comment utiliser le modèle simulé du Télidon. On leur a d'abord montré la solution d'un problème d'exercice, et ensuite on leur a demandé de résoudre eux-mêmes 16 problèmes sans l'intervention de l'expérimentateur. Celui-ci n'est pas intervenu pendant que les participants accomplissaient leur travail.

Pour résoudre chacun des problèmes, le participant devait trouver un renseignement précis dans la base de données. Dans 12 cas sur 16, le participant pouvait trouver la réponse en choisissant des rubriques au hasard dans des pages d'index consécutives jusqu'à ce qu'il parvienne à une feuille de l'arbre. On trouvera au tableau 1 les problèmes qui ont été posés aux participants. Les problèmes étaient formulés de façon à ne contenir aucune référence aux phrases figurant dans les pages d'index. Dans les quatre autres cas, les données n'étaient pas dans la base de données. L'ordre de présentation des problèmes était différent pour chacun des participants et il était conçu de sorte que les effets séquentiels de la présentation des problèmes s'annulent d'un participant à l'autre.

La simulation comportait l'affichage de pages d'index successives sur un terminal vidéo d'ordinateur Volker-Craig 404. Chaque page d'index contenait une liste de phrases dont le nombre pouvait aller jusqu'à neuf. Chacune de ces phrases décrivait les renseignements que contenaient les documents conservés sous cette rubrique. Lorsque le participant choisissait une des rubriques en appuyant sur la touche numérique appropriée du clavier, la page d'index correspondante s'effaçait de l'écran et une autre page était visualisée. La page d'index suivante contenait des rubriques qui constituaient des catégorisations plus précises de renseignements décrits par la rubrique de la page précédente. Une suite de choix de descriptions du renseignement rendait donc les descriptions de plus en plus précises jusqu'à ce que le participant parvienne à la description la plus précise possible. Le choix d'une des rubriques les plus précises produisait un message informant le participant qu'il était parvenu à la solution du problème suivi de la présentation du problème subséquent ou ramenait le participant à la page d'index la plus générale afin qu'il fasse un nouvel essai.

En plus des chiffres figurant sur les pages d'index, le participant pouvait appuyer sur la touche zéro qui le ramenait à la dernière page d'index qu'il avait vue. En appuyant à plusieurs reprises sur la touche zéro, le participant pouvait revenir de plus en plus loin en arrière. S'il oubliait le problème dont il cherchait la solution, il pouvait appuyer sur la touche P pour que celui-ci soit visualisé à nouveau. En appuyant sur la touche S, le participant mettait fin à la recherche et le problème suivant lui était présenté.

TABLEAU 1

Voici la liste des problèmes qui ont été présentés à chacun des participants. Le chiffre qui précède chaque problème indique le nombre de pages à consulter au cours d'une recherche parfaite. Si le chiffre est zéro, les données ne sont pas dans la base de données.

5	Quels sont les derniers résultats dans la Ligue nationale de football?
4	Quel temps fait-il au Manitoba?
4	Trouvez, dans la liste, le deuxième reportage international.
0	Trouvez une recette de poulet frit.
3	Trouvez le nom des ministres du gouvernement de la Saskatchewan.
4	Quel est l'horoscope des Scorpions pour aujourd'hui?
0	Quels sont les derniers numéros gagnants des tirages de Loto Canada?
3	Trouvez un conte pour enfants.
3	Où pouvez-vous emprunter de l'argent?
3	Trouvez des renseignements sur Diefenbaker du temps qu'il était premier ministre.
0	Quelles sont les pharmacies qui seront ouvertes dimanche prochain?
2	Trouvez le numéro de téléphone de la police.
4	Comment les pages du Télidon sont-elles numérotées?
4	Jouer à "Patrouille du cosmos".
4	Trouvez une liste d'expositions d'artisanat.
0	Trouvez une carte de la Suisse.

TABLEAU II

Voici les quatre pages dans lesquelles les participants ont commis le plus d'erreurs. La première page est située à la racine de l'arbre et les trois autres sont celles qu'on obtient en choisissant les rubriques 1, 3 et 6 de la première page. Le premier chiffre qui précède chaque article indique le nombre total d'erreurs qu'ont commises les participants lorsqu'ils ont choisi cet article.

12	1	ACTIVITÉS PERSONNELLES
1	2	AFFAIRES
1	3	GOUVERNEMENT
5	4	NOUVELLES, MÉTÉO ET SPORTS
1	5	URGENCE
1	6	POUR COMPRENDRE LE TÉLIDON
3	1	OU ALLER POUR SE DIVERTIR
0	2	NOTES PERSONNELLES
0	3	VOYAGES
3	4	AIDE ET CONSEILS PERSONNELS
0	5	ACTIVITÉS POUR LE FOYER
0	6	CALENDRIER DES ÉVÉNEMENTS
0	7	ACTIVITÉS ÉDUCATIVES
0	8	TABLES ET STATISTIQUES
0	1	LA CHAMBRE DES COMMUNES
0	2	LE SÉNAT
0	3	LA JUSTICE
0	4	RAPPORTS DU GOUVERNEMENT FÉDÉRAL
7	5	PREMIERS MINISTRES CANADIENS
0	6	D'AUTRES PREMIERS MINISTRES
0	7	GOUVERNEMENTS PROVINCIAUX
0	8	ADMINISTRATIONS MUNICIPALES
4	1	COMMENT SE SERVIR DU TÉLIDON
0	2	EXPLICATIONS SUR LE TÉLIDON
7	3	INDEX DES PAGES DU TÉLIDON

CHAPITRE 2

UTILISATION D'UN INDEX RAMIFIÉ CONTENANT TROIS DÉFAUTS DE CONCEPTION

Thomas Whalen et Candy Mason

UTILISATION D'UN INDEX RAMIFIÉ CONTENANT TROIS DÉFAUTS DE CONCEPTION

RÉSUMÉ

Dans la présente étude, nous cherchons à déterminer de quelle façon différents types de défauts de conception du système d'index ramifié influent sur la recherche documentaire. Les fournisseurs d'informations, connaissant l'importance relative de ces défauts, seront en mesure de mieux concevoir un index pour systèmes vidéotex.

Nous avons introduit dans un système d'index ramifié ressemblant beaucoup à celui de la base de données de démonstration du système vidéotex Télidon quatre cas de chacun des trois types de défauts de conception. Le premier type consistait en une catégorisation de l'information sous une mauvaise rubrique. Le deuxième type de défauts était l'inclusion de deux rubriques synonymes dans la même page. Le troisième type consistait à inclure dans une page une rubrique ambiguë, qui n'apportait aucun renseignement utile à l'utilisateur.

Nous avons demandé à chacun des 21 fonctionnaires qui ont participé à l'expérience de trouver 16 éléments d'information différents au moyen d'un programme d'ordinateur qui simulait un système d'index ramifié du Télidon. Nous avons comparé plusieurs mesures de l'efficacité des recherches effectuées dans cet index défectueux à une étude précédente. Les mêmes tendances sont apparues. Tous les défauts réduisaient l'efficacité de la recherche, mais les rubriques synonymes ou ambiguës avaient des effets beaucoup moins importants que ceux des mauvaises catégorisations.

Ces résultats nous ont incités à faire les recommandations qui suivent. Il faut éviter les trois types de défauts, chacun ayant un effet négatif sur chacune des mesures de l'efficacité de la recherche utilisées dans la présente étude. Cependant, lorsqu'un défaut est inévitable, un titre plus général est préférable à un titre trop précis parce que les rubriques synonymes ou ambiguës qui en résulteront auront, sur la recherche documentaire, des effets moins nuisibles que les catégorisations erronées. Dans tous les cas, un système d'index doit être soumis à des essais empiriques pour détecter la présence de défauts. En effet, l'indexateur pourrait ne pas remarquer ces défauts, ce qui rendrait la recherche plus difficile.

INTRODUCTION

La présente expérience a pour objet de mesurer la probabilité selon laquelle la recherche documentaire sera moins fructueuse et les temps de recherche plus longs en raison des défauts de conception du système d'index d'une base de données.

Le système d'index à l'étude consiste en une série de pages de menu agencées en structure hiérarchique. Ce type d'index, souvent appelé index à structure ramifiée, a été proposé comme le meilleur système de recherche documentaire pour le non-initié, c'est-à-dire l'utilisateur qui se sert rarement d'une base de données, et donc comme celui qui convient le mieux au système Télidon et aux autres systèmes vidéotex.

Dans toute base de données avec index, l'efficacité de la recherche documentaire dépend de la conception de ce dernier. Toute partie du système d'index qui cause des difficultés à l'utilisateur est par définition un défaut, même si le concepteur perçoit ce système comme logique ou sensé. Pareille situation risque de se produire parce que la sémantique des langues naturelles n'est pas fixe, mais varie d'un individu à l'autre. Ainsi, deux personnes peuvent diverger d'opinion sur l'opportunité d'un nom pour désigner une catégorie (Rosch, 1975).

La possibilité que certaines parties d'un système d'index soient en principe valables mais qu'elles contiennent néanmoins des défauts selon ce critère peut rendre très difficile la conception de systèmes d'index. Le seul moyen de savoir si une partie donnée d'un système d'index contient des défauts pour la plupart des utilisateurs, c'est de vérifier si un groupe sélectionné d'utilisateurs réussit à trouver des informations en se servant de cette partie de l'index.

Dans des études précédentes (Lee et Latrémouille, 1979; Whalen et Latrémouille, 1980), on a conclu que le système d'index était déficient parce que les utilisateurs éprouvaient de la difficulté à trouver l'information voulue. On a alors émis des hypothèses quant à la nature des défauts du système, mais sans pouvoir les confirmer. Par exemple, il était impossible de dire quels types de défauts étaient le plus nuisible à la recherche documentaire. La seule façon de s'assurer que la plus grande partie d'un effet observé résulte d'un défaut de conception donné est d'introduire de façon délibérée, dans la base de données, un défaut plus grave que les autres types de défauts secondaires. La présente étude examine donc, dans un cadre expérimental, la gravité de défauts de conception introduits volontairement dans l'index d'une base de données.

L'expérience a porté sur trois types de défauts. On parle de catégorisations erronées lorsqu'une rubrique d'une page d'index donne accès à une deuxième page d'index contenant une rubrique qui devrait, au sens de l'utilisateur, figurer sous un autre nom, dans la première page d'index. Par exemple, si la rubrique "CONSOMMATION" renvoie à une autre page d'index qui ne contient pas de rubriques sur les expositions d'artisanat, mais que la rubrique "MÉTHODES ET SERVICES DE PUBLICITÉ ET DE COMMERCIALISATION" conduit à une page contenant une rubrique "EXPOSITIONS D'ARTISANAT", nous dirons que la rubrique "EXPOSITIONS D'ARTISANAT" est mal catégorisée.

Les rubriques synonymes constituent le deuxième type de défaut de conception. Si un utilisateur ne voit pas de différence entre deux rubriques figurant sur une même page d'index, il n'aura pas de motif d'en choisir une plutôt que l'autre. Par exemple, si les rubriques "DISTRACTIONS" et "DIVERTISSEMENTS" apparaissent sur la même page, la plupart des gens vont les considérer comme synonymes.

Le troisième type de défaut consiste en des titres ambigus. Si, en voyant le titre d'une rubrique l'utilisateur ne voit pas clairement de quoi il s'agit, c'est que ce titre est ambigu. Ce serait le cas de "RENSEIGNEMENTS DIVERS", parce que ce titre n'indique pas clairement le type de renseignement auquel il renvoie. Il importe de remarquer qu'il ne s'agit pas ici d'une rubrique dont la signification est différente de celle qu'avait prévue l'utilisateur. Dans ce cas, l'utilisateur comprend le sens de la rubrique, mais toutes les données classées sous celle-ci sont dans la mauvaise catégorie.

Il est évident que ces trois types de défaut ne s'excluent pas l'un l'autre. Par exemple, il est possible que les rubriques "RENSEIGNEMENTS DIVERS" et "INFORMATION GÉNÉRALE" apparaissant sur la même page soient considérées à la fois comme rubriques synonymes et comme titres ambigus.

Pour les fins de notre étude, nous avons volontairement introduit quatre défauts de chaque type dans une structure d'index qui avait déjà été examinée par Whalen et Latrémouille (1980). Cet index modifié a ensuite été mis à l'épreuve comme précédemment. Les problèmes présentés aux participants et tous les autres aspects de l'étude étaient les mêmes. Seul le nombre des sujets était différent. De cette façon, il était plus facile de comparer les résultats des différentes recherches à ceux de l'étude précédente.

Aussi bien dans cette expérience que dans la précédente, nous avons informé les sujets que certains renseignements demandés ne se trouvaient pas dans la base de données. Ainsi, les sujets pouvaient mettre fin à leur recherche s'ils ne pouvaient trouver la réponse. Dans l'étude précédente, nous avons constaté que la probabilité qu'une erreur soit commise lors d'une recherche documentaire était forte. Cependant, lorsque l'information recherchée se trouvait effectivement dans la base de données, la plupart des participants ont mené leur recherche à bonne fin, plutôt que de conclure que le problème était insoluble.

MÉTHODE

PARTICIPANTS

Vingt et un employés du ministère des Communications ont participé individuellement à des séances expérimentales de recherche d'une durée approximative d'une heure.

MÉTHODE

Tout d'abord, chaque participant a reçu des instructions sur l'utilisation simulée du Télidon. Ensuite, après un exercice, chacun a eu à résoudre 16 problèmes expérimentaux.

Les problèmes présentés aux sujets étaient les mêmes que ceux qui avaient servi dans l'expérience précédente de Whalen et Latrémouille (1980). Les participants devaient retrouver un élément d'information particulier dans une base de données, mais les problèmes étaient formulés de façon à ne contenir aucun des termes utilisés dans les pages d'index, sauf dans une page du niveau inférieur. La consultation de la base de données permettait de trouver réponse à 12 des 16 problèmes présentés. On avait indiqué aux participants que certains problèmes étaient sans réponse, mais sans préciser lesquels, ni le nombre. Les problèmes ont été présentés dans un des quatre ordres différents possibles de façon à réduire au minimum les effets de la présentation. Les 16 problèmes sont reproduits intégralement à l'Annexe A.

Aux fins de la simulation, un terminal Volker-Craig 404 a été utilisé pour afficher successivement les pages d'index. Chaque page contenait au plus neuf rubriques numérotées. Exception faite des

modifications requises pour introduire les quatre défauts de chaque type, les pages étaient celles que l'on retrouvait dans la base de données de démonstration du Télidon en novembre 1979 et qui ont aussi servi dans l'étude précédente de Whalen et Latrémouille (1980). L'emplacement exact de chaque défaut est indiqué à l'Annexe B.

Lorsque le participant choisissait une rubrique en appuyant sur le numéro correspondant du clavier du terminal, une autre page de rubriques du niveau suivant apparaissait à l'écran. Outre les touches numériques correspondant aux numéros des rubriques, le clavier comportait la touche "0" qui permettait au participant de revenir à la page d'index précédente. Celui-ci revenait au tronc de l'arbre en appuyant plusieurs fois de suite sur "0" ou lorsqu'il choisissait une rubrique dans une extrémité des ramifications. Le participant pouvait rappeler la question à l'écran et mettre fin à la recherche en appuyant respectivement sur "P" et "S".

RÉSULTATS

Les résultats de la présente expérience sont résumés et comparés à ceux de l'expérience précédente menée par Whalen et Latrémouille (1980) au tableau 1.

RENDEMENT GLOBAL

Les participants ont trouvé une solution à 150 des 252 problèmes solubles. Pour ce qui est des 84 autres problèmes, la solution ne se trouvait pas dans la base de données et les participants ont correctement mis fin à leurs recherches. Ils ont donc correctement interrompu leurs recherches ou trouvé une solution pour 234 des 336 problèmes, soit un taux de réussite de 70%. Ils ont interrompu prématurément leurs recherches dans 102 des 252 cas pour lesquels il existait une solution, soit dans 40% des cas. Ces résultats diffèrent de ceux obtenus dans une étude précédente de Whalen et Latrémouille (1980) où la probabilité d'une fin prématurée était de 13,5%. Nous en avons donc déduit que les utilisateurs sont beaucoup plus portés à mettre fin prématurément à leurs recherches lorsque des défauts sont introduits dans la base de données.

Dans la présente expérience, les participants ont fait le bon choix dans 60% des cas où la réponse se trouvait dans la base de données. Le risque moyen d'erreur par page s'élevait donc à 40%. Dans deux expériences précédentes (Whalen et Latrémouille, 1980; Lee et Latrémouille, 1979), la probabilité moyenne d'erreur par page était d'environ 15%. Ainsi, dans l'ensemble, il est beaucoup plus probable qu'un utilisateur fera des choix erronés si la structure de l'index comporte des défauts.

TABLEAU 1

Résumé comparatif des résultats des recherches documentaires faites dans la présente expérience et dans celle de Whalen et Latrémouille (1980).

	A	B	C	D
Résultats d'ensemble:				
présente expérience	40.0%	40.0%	10.6 pages	58.5%
expérience précédente	13.5	15.0	8.3	24.2
Catégorisations erronées:				
présente expérience	95.2	0.01	14.0	94.3
expérience précédente	16.6	-	4.4	28.7
Rubriques synonymes:				
présente expérience	17.8	45.0	7.0	9.7
expérience précédente	8.4	-	5.7	7.3
Titres ambigus:				
présente expérience	14.2	44.0	3.7	17.2
expérience précédente	12.5	-	3.0	8.7

Colonne A: Probabilité qu'une recherche soit interrompue prématurément.

Colonne B: Probabilité d'un choix erroné dans une page.

Colonne C: Nombre de pages consultées pour trouver la réponse.

Colonne D: Pourcentage du total des pages consultées au cours de recherches interrompues prématurément.

Nous avons déterminé les temps de recherche d'après le nombre de fois qu'un participant choisissait une rubrique conduisant à une autre page d'index. Pour ce qui des problèmes résolubles, les participants ont consulté en moyenne 10,6 pages. Des recherches sans aucune erreur auraient en fait nécessité l'extraction de 3,58 pages en moyenne seulement.

Étant donné que les participants à cette expérience faisaient pour la première fois une recherche dans la base de données, ils ont consulté un moins grand nombre de pages après s'être familiarisés avec la tâche et la base de données. Chaque participant a consulté en moyenne 14 pages pour le premier problème qu'il a réussi à résoudre. Pour chacun des autres problèmes résolus, les participants n'ont consulté en moyenne que 5,65 pages. Autrement dit, après une première recherche fructueuse, les participants en venaient à trouver plus rapidement les renseignements voulus dans la base de données.

Le nombre moyen de pages consultées par les participants pour leur première recherche interrompue prématurément est de 14,81. Cependant, pour les deuxième, troisième et quatrième problèmes pour lesquels la recherche a été inachevée, cette moyenne n'était plus que de 10,24 pages, 9,6 pages et 6,9 pages respectivement.

Dans le cas des problèmes dont la recherche a été correctement interrompue parce que la solution était introuvable, les participants ont tout d'abord consulté 16,81 pages en moyenne, mais seulement 11,81 pages en moyenne lors de leurs recherches ultérieures. Ainsi, tant pour les recherches interrompues que pour les problèmes résolus, les participants en sont venus à consulter un nombre plus restreint de pages une fois qu'ils s'étaient familiarisés avec la base de données.

RENDEMENT DES PARTICIPANTS RELATIVEMENT AUX DÉFAUTS DE CONCEPTION

Lorsque la solution d'un problème se retrouvait sous un titre inexact, la probabilité que les participants fassent le bon choix était inférieure à 1%. Lorsque les participants devaient choisir entre deux rubriques synonymes, ils choisissaient la bonne rubrique selon une probabilité de 45%. Lorsque la réponse se trouvait sous la première rubrique synonyme, la probabilité n'était plus que de 37%. Lorsque la page contenait un titre ambigu qui ne devait pas être choisi dans un cheminement exact, la probabilité que les sujets choisissent la bonne rubrique était de 44%.

Comme prévu, la catégorisation erronée des données est le défaut de conception le plus important. Lorsqu'un renseignement se retrouve sous un titre qui n'est vraiment pas représentatif, il est bien peu probable que l'utilisateur le choisira. Aussi, comme nous l'avions prévu, la probabilité que l'utilisateur fasse un choix exact entre deux rubriques synonymes est inférieure à 50%. Elle serait exactement de 50% si les sujets choisissaient toujours la même rubrique, mais, quelquefois, aucune des rubriques synonymes n'est choisie. Il est intéressant de noter que les participants choisissent la deuxième rubrique synonyme plus souvent que la première. Cette constation nous porte à croire que l'utilisateur n'a pas interrompu sa recherche à la première rubrique qui lui semblait contenir la solution, mais qu'il a poursuivi sa lecture jusqu'à ce qu'il rencontre au moins une autre rubrique qui lui semblait convenir.

COMPARAISON AVEC LES DONNÉES ANTÉRIEURES

En vue d'évaluer de façon plus complète l'effet des défauts de conception de l'index sur l'efficacité de la recherche, nous avons comparé les recherches faites dans le cadre de la présente expérience

à celles qui avaient été faites dans un index dans lequel aucune erreur n'y avait été insérée, dans une expérience précédente (Whalen et Latrémouille, 1980). Dans la présente étude, trois réponses étaient classées sous la mauvaise rubrique, six se trouvaient sous une rubrique synonyme et deux sous un titre ambigu.

Dans le cas des problèmes dont la solution ne se retrouvait pas sous la bonne rubrique, les participants à la présente expérience ont mis fin prématurément à leurs recherches dans 60 des 63 cas possibles. La probabilité de mettre fin à une recherche sans trouver la réponse est donc de 95,2%. Pour les mêmes problèmes, les participants à l'expérience précédente n'ont interrompu prématurément leurs recherches que dans 4 cas sur 24. La probabilité d'interrompre une recherche avant de trouver la réponse était dans ce cas de 16,6% seulement. Lorsqu'une catégorisation erronée était introduite dans le cheminement de la recherche, la probabilité de mettre fin à la recherche sans trouver la solution passait à 78,6%.

Pour ce qui est des problèmes dont la solution se trouvait sous une des deux rubriques synonymes, les participants ont interrompu prématurément leurs recherches dans 15 des 84 cas possibles, soit 17,8%. Dans l'expérience précédente, lorsque la rubrique sous laquelle était classée la réponse était clairement formulée, seulement trois recherches sur 32 ont été interrompues prématurément, soit 9,4%. Ainsi, la probabilité de mettre fin prématurément à une recherche s'élève à 8.4% de plus lorsque des rubriques synonymes sont introduites dans le cheminement de la recherche.

Quant aux problèmes pour lesquels il fallait chercher des renseignements classés sous des rubriques ambiguës, les participants ont prématurément interrompu leurs recherches dans 6 des 42 cas possibles, soit 14,2%. Dans l'expérience précédente, deux des 16 recherches ont été interrompues prématurément, soit 12,5%. Ainsi, la probabilité de mettre fin prématurément à une recherche s'élevait à

1,7% de plus lorsque des titres ambigus étaient introduits dans le cheminement de la recherche.

Dans le cas des problèmes résolus, le nombre de pages consultées pour trouver la réponse a servi à déterminer le degré de difficulté de la recherche. Dans la présente expérience, les participants ont consulté en moyenne 14 pages avant de résoudre les problèmes dont la réponse avait été classée dans une mauvaise catégorie. Dans l'expérience précédente, pour les mêmes problèmes, les participants avaient consulté en moyenne 4,4 pages. Lorsqu'une catégorisation erronée est introduite dans l'index, les utilisateurs doivent donc consulter en moyenne 9,6 pages de plus pour trouver la réponse, ce qui représente une augmentation de 122,7% du temps de recherche par rapport à l'expérience précédente.

Lorsque les renseignements recherchés étaient classés sous une rubrique synonyme, les participants ont consulté en moyenne 6,99 pages pour trouver la réponse aux questions, alors qu'ils en ont consulté en moyenne 5,72 dans l'expérience précédente. Autrement dit, la durée de la recherche a subi une augmentation de 22,2% dans la présente expérience.

Lorsque les renseignements recherchés étaient classés sous une rubrique ambiguë, les participants ont consulté en moyenne 3,75 pages pour trouver la réponse, alors qu'ils en ont consulté, en moyenne, 3 dans l'expérience précédente. La durée de la recherche a donc augmenté de 25% par rapport à l'expérience précédente.

EFFICACITÉ GLOBALE DE LA BASE DE DONNÉES

Nous avons pris comme mesure générale de l'efficacité de la base de données la proportion du nombre total de pages consultées pour

lesquelles il y a eu interruption prématurée de la recherche. Ces efforts auraient dû conduire à une recherche fructueuse, mais cela n'a pas été le cas. Dans l'étude précédente, 24,2% des pages consultées pour trouver des renseignements présents dans la base de données ont été consultées au cours de recherches qui ont été interrompues prématurément. Ce pourcentage est passé à 58,5% lorsque des défauts de conception ont été introduits dans la base de données. Ainsi, c'est beaucoup plus que la moitié de toutes les pages consultées pour trouver des renseignements existants qui ont donné lieu à des recherches infructueuses lorsque la base de données comportait des défauts.

Si l'on considère uniquement les problèmes dont la solution était classée sous une mauvaise catégorie, cette proportion passe de la première à la deuxième expérience de 28,7% à 94,3%. La proportion passe de 7,26% à 9,73% dans le cas des rubriques synonymes et de 8,7% à 17,17% dans le cas des titres ambigus. Ainsi, une fois encore, les catégorisations erronées sont de loin le défaut de conception qui cause le plus d'effet.

DISCUSSION

Les résultats obtenus montrent une tendance très constante. Tout type de défaut de conception de l'index a pour effet de réduire l'efficacité de la recherche selon toute mesure raisonnable que l'on utilise. Ceci n'est pas surprenant si l'on considère l'ampleur des défauts introduits dans la base de données.

Cependant, les défauts de conception n'ont pas tous les mêmes effets nuisibles. Les catégorisations erronées causent un plus grand nombre d'erreurs puisqu'elles empêchent pratiquement de trouver la solution d'un problème. D'autre part, il est étonnant de constater que les rubriques synonymes ou les rubriques ambiguës ont des effets bénins. Elles se traduisent par un temps de recherche un peu plus long et augmentent la probabilité que les utilisateurs ne trouvent pas la solution recherchée, mais leurs effets sont très minimes, surtout si l'on tient compte de la gravité de tous les défauts de conception introduits dans la base de données.

Il existe une relation entre les types de défauts de conception examinés dans la présente étude. Une fois les données groupées dans une catégorie, le choix d'un titre décrivant cette catégorie aura tendance à produire des catégorisations erronées si le titre est trop précis, alors que ce titre sera ambigu et aura tendance à en chevaucher d'autres s'il est trop général.

Compte tenu de ces résultats, nous faisons les recommandations suivantes:

1) Il faut grouper les données dans des catégories aussi distinctes et logiques que possible sinon un des défauts susmentionnés sera inévitable et la présence de n'importe quel de ces défauts réduira l'efficacité de la structure d'index.

2) Après avoir établi une catégorie, il faut lui donner le titre le décrivant le plus exactement possible. En effet, comme nous l'avons déjà dit, une description inexacte donnera lieu à un défaut.

3) Au moment de choisir un nom de catégorie, s'il est impossible de trouver un titre parfait, il faut s'assurer que le titre choisi regroupe bien toutes les données auxquelles il renvoie. Le titre risque d'être plus ambigu ou de recouper celui qui renvoie à une autre catégorie, mais ces défauts causent un plus petit nombre d'erreurs que les catégorisations erronées.

4) Lorsqu'une structure d'index a été établie, on doit la soumettre à des essais empiriques pour déterminer la gravité des défauts qui peuvent y avoir été introduits. Les résultats recueillis montrent qu'un défaut de conception de l'index peut avoir des effets sérieux. Il faudra probablement reprendre plusieurs fois ces étapes avant d'en arriver à un système d'index efficace.

Ces recommandations valent seulement pour la conception des systèmes d'index par choix de menus qui permettent à l'utilisateur de revenir à la page de menu précédente et de modifier ainsi facilement ses décisions erronées s'il découvre qu'il fait fausse route. Dans les autres systèmes comme les systèmes de recherche par mots repères, où le choix d'un terme trop général peut se traduire par une surabondance de données, l'importance relative de ces types de défauts de conception peut être très différente.

Le champ d'application de ces recommandations peut aussi être limité du fait qu'elles sont fondées sur les cas les plus graves possibles de chaque type de défaut. Comme il n'y a pas d'élément empirique prouvant le contraire, nous devons supposer que l'importance relative de chaque type de défaut sera la même pour des cas moins graves. Cette supposition semble justifiée par l'ampleur des effets constatés, mais d'autres recherches empiriques l'infirmeront peut-être.

OUVRAGES DE RÉFÉRENCE

Lee, E.A. et Latrémouille, S., Évaluation de la structure arborescente de l'information dans la base de données du Télidon. (Note technique BRG 79-12). Ministère des Communications, Ottawa. Décembre 1979. Aussi disponible dans Recherche en comportement sur Télidon 1. Ministère des Communications, Ottawa, 1980, pp. 228 à 236.

Rosch, E., Cognitive representations of semantic categories. Journal of Experimental Psychology: General. 1975, 104, 3, pp.192 à 233.

Whalen, T.E. et Latrémouille, S., L'efficacité d'un index ramifié lorsque l'existence de données est incertaine. (Note technique nº BRIC 80-3). Ministère des Communications, Ottawa. 1980.

ANNEXE A

Voici les problèmes qui ont été présentés à chaque sujet. Les chiffres qui figurent entre parenthèses indiquent le nombre minimal de pages qui doivent être consultées pour trouver la réponse voulue. Un zéro signifie que l'information recherchée ne se trouve pas dans la base de données.

1. Quels sont les derniers résultats dans la Ligue nationale de football? (5)

2. Quel temps fait-il au Manitoba? (4)

3. Trouvez, dans la liste, le deuxième reportage international. (4)

4. Trouvez une recette de poulet frit. (0)

5. Trouvez le nom des ministres du gouvernement de la Saskatchewan. (3)

6. Quel est l'horoscope des Scorpions pour aujourd'hui? (4)

7. Quels sont les derniers numéros gagnants des tirages de Loto Canada? (0)

8. Trouvez un conte pour enfant. (3)

9. Où pouvez-vous emprunter de l'argent? (3)

10. Quand Diefenbaker a-t-il été premier ministre? (3)

11. Quelles sont les pharmacies qui seront ouvertes dimanche prochain? (0)

12. Trouvez le numéro de téléphone de la police. (2)

13. Comment les pages du Télidon sont-elles numérotées? (4)

14. Jouez à "Patrouille du cosmos". (4)

15. Trouvez une liste d'expositions d'artisanat. (4)

16. Trouvez une carte de la Suisse. (0)

ANNEXE B

Voici la liste de chacun des défauts de conception introduits dans l'index.

I. CATÉGORISATIONS ERRONÉES

1. "Résultats sportifs" a été déplacé de "Nouvelles, météo et sports" à "Urgence".

2. "Cartes" a été déplacé de "Voyages" à "Assistance et conseils personnels".

3. "Expositions d'artisanat" a été déplacé de "Consommation" à "Méthodes et services de publicité et de commercialisation".

4. "Ministres de la Saskatchewan" a été déplacé de "Gouvernement provincial" à "Chambre des Communes".

II. RUBRIQUES SYNONYMIQUES

1. "Gouvernement" et "Industrie" ont été regroupés sous le titre "Gouvernement et industrie".

2. "Distractions" et "Divertissements" figuraient sur la même page.

3. La rubrique des nouvelles locales comprenait une "Première section" et une "Deuxième section".

4. "Détente à la maison" et "Détente chez soi" figuraient sur la même page.

III. TITRES AMBIGUS

1. Le titre "Renseignements divers" apparaissait sur une page.

2. Le titre "Autres renseignements" apparaissait sur une page.

3. Le titre "Autres reportages" apparaissait sur une page.

4. Le titre "Urgences diverses" apparaissait sur une page.

CHAPITRE 3

ÉTUDE DE L'EFFICACITÉ D'UTILISATION D'UN SYSTÈME DE RECHERCHE DOCUMENTAIRE TÉLIDON

Scott A. McEwen

ÉTUDE DE L'EFFICACITÉ D'UTILISATION D'UN SYSTÈME DE RECHERCHE DOCUMENTAIRE TÉLIDON[1]

RÉSUMÉ

Le développement des systèmes informatisés de recherche documentaire a suscité un énorme intérêt pour ces systèmes et donné lieu à beaucoup de recherches sur l'interaction homme-machine. La présente expérience, qui est la première expérience de consultation d'une base de données Télidon réelle, avait pour but de vérifier l'efficacité d'utilisation du système. On a demandé à 24 utilisateurs non initiés de trouver dans la base de données les réponses à 16 questions. Le temps de recherche et le nombre d'erreurs commises par les participants constituaient les principales variables dépendantes. Les résultats de l'étude ont démontré que les utilisateurs éprouvaient certaines difficultés à repérer l'information dans une base de données à structure hiérarchique. L'expérience a également révélé que certaines parties de la structure arborescente posaient des problèmes. Ces résultats sont en accord avec ceux d'études précédentes. La conclusion correspond également à celle qui découle des études antérieures. Les gens peuvent éprouver des difficultés lorsqu'ils consultent des bases de données à structure hiérarchique semblables à celle qui a fait l'objet de l'expérience. A la lumière de ces données, il est possible de déterminer où se produisent les erreurs dans une structure arborescente et de proposer les améliorations appropriées.

1 L'auteur tient à remercier Eric Lee, Paul Muter, Thomas Whalen et Susane Latrémouille, membres du Groupe de recherche en comportement du ministère des Communications, de leurs commentaires et de leur collaboration.

INTRODUCTION

Au cours des dernières années, on a mis au point un certain nombre de systèmes vidéotex de recherche documentaire. Parmi les systèmes les plus importants et les plus connus, il y a le Télidon, mis au point par le Centre de recherche du ministère des Communications à Ottawa, le Prestel britannique et l'Antiope français. Ces systèmes suscitent de plus en plus d'intérêt et de recherche du fait que bientôt ils seront utilisés dans les foyers.

Le développement futur de ces systèmes repose non seulement sur les réalisations techniques récentes qui ont permis leur construction mais aussi sur les facteurs humains propres à l'interaction homme-machine. L'accueil que réserve le public à ces systèmes sera fonction de leur compatibilité avec la démarche cognitive et le comportement réactionnel humains. Si les gens éprouvent des difficultés à utiliser un système donné, ils ne seront évidemment pas portés à l'utiliser. C'est pourquoi, avant de commercialiser un produit, il est indispensable d'entreprendre des recherches afin de déceler les éventuels problèmes qu'il pourrait poser à l'utilisateur.

La présente expérience a été conçue pour examiner plusieurs questions reliées à l'interaction homme-machine lors de l'utilisation d'une base de données Télidon. La question fondamentale consiste à savoir si le Télidon est un système facile à utiliser. Pour y répondre nous avons demandé à un groupe de personnes de chercher des éléments d'information dans une base de données Télidon. Nous avons opté pour la méthode qui se rapproche le plus de l'utilisation réelle du Télidon.

Une expérience de ce genre permet non seulement de déterminer dans quelle mesure le Télidon est facile à utiliser mais aussi d'examiner la structure arborescente hiérarchique de certaines bases

de données Télidon. S'agit-il de la meilleure structure d'agencement de l'information? Quels problèmes rencontrent les gens lorsqu'ils consultent les bases de données? Quelles modifications peut-on apporter à la structure arborescente afin de minimiser ou éliminer les problèmes?

Ces questions et bien d'autres ont déjà été abordées dans certaines études portant sur les bases de données à structure hiérarchique. En 1979, Van Ness et Tromp ont fait une étude exploratoire sur l'utilisation d'un système vidéotex (Prestel) de faibles dimensions afin de recueillir de l'information sur les facteurs humains rattachés au vidéotex. Ils ont demandé à sept personnes, qui n'avaient jamais utilisé le vidéotex, d'effectuer sept opérations de recherche dans une base de données comportant 75 pages d'index et 900 pages de données (document). Un expérimentateur a chronométré la durée de chaque épreuve de recherche indiquant parfois aux utilisateurs que l'information demandée se trouvait effectivement dans la base de données.

Les résultats de cette étude ont démontré que, pour l'ensemble des épreuves, les participants ont consulté deux fois plus de pages que le nombre minimum nécessaire pour trouver l'information. Le temps de recherche moyen par épreuve, extrapolé à partir des données de l'étude, s'établit à environ 2,7 minutes (162 secondes). Le temps de recherche s'échelonnait de 1,7 minutes (90 secondes) à 16 minutes (960 secondes) environ. A la lumière de ces résultats, Van Ness et Tromp mettent en doute la facilité d'utilisation du vidéotex.

Lors d'une étude menée en 1979 par Lee et Latrémouille, on a demandé aux participants de choisir des rubriques dans des menus consécutifs (pages d'index) présentés sur des fiches de 13 x 18cm afin de repérer l'information demandée. Cette expérience avait pour but d'évaluer, d'une part, la structure arborescente hiérarchique de certaines bases de données Télidon et, d'autre part, l'efficacité d'utilisation d'un système Télidon réel. Les utilisateurs savaient d'avance que les renseignements à trouver se trouvaient dans la base de données. Pendant l'expérience, on leur faisait savoir immédiatement si leurs choix étaient bons ou mauvais. Lorsque les utilisateurs se trompaient, on recueillait leurs commentaires afin de déterminer les causes des erreurs.

Cette expérience a démontré que l'arbre présentait certaines déficiences et que les gens éprouvaient des difficultés à repérer l'information. Par exemple, la probabilité d'erreur pour chaque choix donné a été établie à 0,14. On a également remarqué que 53% des erreurs se sont produites aux deux premiers niveaux de la structure arborescente et que chaque personne a commis des erreurs dans environ la moitié des cas. La conclusion de cette étude laisse entendre que les gens, en général, auront des difficultés à repérer l'information dans les bases de données hiérarchisées semblables à celle utilisée dans l'expérience.

Pour faire suite à l'étude décrite ci-dessus, Whalen et Latrémouille ont mené, en 1980, une autre expérience. Cette fois, au lieu de fiches, on a utilisé un gros ordinateur en temps partagé pour simuler le système et la base de données Télidon. La façon de procéder a été quelque peu modifiée par Whalen et Latrémouille afin que la situation s'apparente davantage à celle qui existe lorsqu'on utilise un système Télidon réel. Contrairement à la méthode utilisée au cours des études précédentes, on a informé les utilisateurs que certains éléments d'information demandés ne se trouvaient pas dans la base de données. Lorsque les gens utilisent un système Télidon réel,

on ne sait pas si l'information recherchée se trouve vraiment dans la base de données. Par conséquent, lors de l'expérience, les participants avaient le loisir d'arrêter une recherche donnée à tout moment et de choisir un nouveau problème. Lorsqu'on consulte le Télidon, on ne sait pas non plus si les choix qu'on fait permettront de trouver l'information désirée. Par conséquent, les participants n'ont pas été informés pendant l'expérience, de l'exactitude ou de l'inexactitude de leurs choix.

L'étude menée par Whalen et Latrémouille a révélé encore une fois des problèmes dans la structure arborescente et le rendement de l'utilisateur. Par exemple, la probabilité d'erreur pour chaque choix donné était de 0,15. Environ 80% de toutes les erreurs se sont produites dans les quatre pages des deux premiers niveaux de l'arbre. Les conclusions de la présente étude sont semblables à celles de l'étude précédente menée par Lee et Latrémouille, selon lesquelles les utilisateurs du Télidon risquent fort de se tromper lorsqu'ils cherchent de l'information dans les pages d'index disposées en structure arborescente.

L'importance de la présente étude réside dans le fait qu'il s'agit d'une première expérience de recherches menée à l'aide d'un vrai système Télidon. Afin de combler les lacunes de la démarche des études précédentes de Van Ness et Tromp et de Lee et Latrémouille, on a, d'une part, informé les participants que certains éléments d'information demandés pouvaient ne pas se trouver dans la base de données et, d'autre part, omis volontairement de leur indiquer si leurs choix étaient bons ou mauvais. Le temps de recherche et le nombre d'erreurs commises constituaient les deux principales mesures dépendantes. La présente étude est aussi la première à mesurer le temps de recherche sur un système Télidon. Elle avait pour objet d'évaluer la base de données existante d'un système Télidon interactif (vidéotex) réel, soit la base de données expérimentale du ministère

des Communications de novembre 1979. Un des principaux objectifs était de déterminer combien de temps en moyenne les gens doivent consacrer à une recherche dans un système Télidon (temps de recherche). On était également intéressé à savoir si le type et le nombre d'erreurs commises par les utilisateurs d'un vrai système Télidon seraient les mêmes que dans les études précédentes. L'expérience a été conçue afin de déterminer les secteurs de l'arbre qui présentent des problèmes et de nous permettre de faire des recommandations visant à améliorer la base de données. Un objectif secondaire consistait à obtenir des descriptions verbales des méthodes utilisées par les utilisateurs pour repérer l'information. Un dernier objectif de la présente étude était de stimuler l'intérêt et d'encourager d'autres recherches dans ce nouveau domaine.

MÉTHODE

SUJETS

Vingt-quatre volontaires de la Commission de la Fonction publique du Canada ont participé à cette expérience. Le groupe comprenait 14 hommes et 10 femmes dont la moyenne d'âge était respectivement de 32,7 et de 31,6. Aucun des participants n'avait jamais utilisé le Télidon ou un système électronique de recherche documentaire.

MATÉRIEL

L'expérience s'est déroulée dans une pièce spécialement aménagée pour simuler l'atmosphère d'un "salon". Les utilisateurs étaient assis sur un divan devant un téléviseur placé à environ 2 mètres d'eux. Une ligne téléphonique et un modem assuraient la liaison directe avec la base de données de l'ordinateur. Pour la transmission des demandes de pages à l'ordinateur et la réception des données, on s'est servi d'un terminal Télidon Norpak Mark I relié au modem. Les réponses de l'ordinateur étaient affichées sur un téléviseur couleur modifié de 47,5 centimètres de marque Electrohome. Pour communiquer avec le système, les participants ont utilisé un clavier Télidon dont les dimensions s'apparentent à celles d'une calculatrice de poche. Le temps de recherche a été déterminé à l'aide d'un chronomètre à cristaux liquides Microntер. On a également utilisé un magnétophone pour enregistrer les commentaires de certains participants.

DÉMARCHE

Dans la présente expérience, les 24 participants ont été répartis au hasard entre deux groupes de 12 personnes chacun. Une condition

d'instruction, la verbalisation ou la non verbalisation, a été imposée à chaque groupe. On a montré aux 24 participants comment utiliser le système Télidon en leur indiquant les types de pages et d'information disponibles et en les renseignant sur la façon d'effectuer la recherche (voir annexe A). On a demandé aux participants assujettis à la condition de verbalisation de verbaliser leur procédé de recherche en donnant les raisons ou les stratégies de leurs choix au fur et à mesure qu'ils faisaient les recherches. On a utilisé un magnétophone pour enregistrer les réactions des participants.

La verbalisation a servi à l'obtention de protocoles verbaux destinés à l'établissement d'un modèle de la démarche cognitive dans le processus de recherche. Ce modèle cherchera à expliquer comment les gens cherchent l'information et pourquoi ils font des erreurs. Un tel modèle pourrait nous permettre également d'améliorer les structures arborescentes utilisées pour la recherche documentaire.

(Au moment de rédiger le présent document, les commentaires verbaux des participants n'avaient pas encore été analysés. Une analyse de ces commentaires figurera dans un document ultérieur).

Puisque les utilisateurs n'avaient jamais utilisé le Télidon, on a donné à chacun une démonstration de la façon d'utiliser le clavier du Télidon, ainsi que deux exemples de questions de recherche afin qu'ils sachent exactement ce qu'ils avaient à faire. La séance d'information et la démonstration ont duré environ 20 minutes au total.

Les études et les expériences pilotes antérieures à la présente expérience ont montré qu'on utilise le Télidon de différentes façons. Dans certains cas, on semble chercher des éléments d'information bien spécifiques. Dans d'autres cas, on cherche des éléments d'information de nature générale qui demandent parfois de scruter diverses parties

de la base de données. La consultation au hasard des pages est aussi une activité fréquente. Au cours de la présente expérience, nous avons essayé d'utiliser des tâches représentatives de celles qu'accomplira probablement l'utilisateur type.

La base de données Télidon de novembre 1979, qu'on a utilisée dans la présente expérience, comportait environ 1400 éléments d'information accessibles. Il y avait 900 pages documentaires, c'est-à-dire d'information, et 500 pages d'index. Parmi toutes ces possibilités, on a choisi 16 éléments au hasard. On a ensuite rédigé des questions en rapport avec les données choisies. Ces questions visaient à former un échantillon représentatif des différents types de recherche que l'utilisateur peut faire sur Télidon. (Les questions de recherche figurent à l'annexe B). Il y avait trois différents types de questions. Parmi les 16, dix étaient de nature "spécifique" puisqu'il n'y avait qu'une seule réponse pour chaque question dans la base de données. Trois questions étaient d'ordre "général", c'est-à-dire que l'on pouvait trouver la réponse à plus d'un endroit dans l'arbre. Les trois autres étaient des questions "sans réponse" puisque la base de données ne contient aucune donnée se rapportant directement à ces questions. En suivant un cheminement particulier, les utilisateurs pouvaient déterminer avec précision si la base de données contenait les réponses à ces trois questions. Ces questions étaient donc du type spécifique sauf que l'information ne se trouvait pas dans la base de données.

Les questions sans réponse avaient pour but de mieux simuler l'utilisation du Télidon puisque les utilisateurs ne trouveront pas réponse à toutes leurs questions sur un système commercial Télidon. Ces questions permettent également d'observer si l'utilisateur réagit à ce type de question de la même façon qu'il le fait dans le cas des questions dont les réponses se trouvent dans la base de données.

On a indiqué aux participants que les réponses à certaines questions ne se trouvaient pas dans la base de données. On leur a demandé d'informer l'expérimentateur s'ils croyaient que la réponse à une question donnée ne se trouvait pas effectivement dans la base de données. C'est l'utilisateur qui devait décider si l'information s'y trouvait ou non. Les recherches pouvaient donc se terminer de deux façons. Lorsque l'utilisateur trouvait la bonne réponse, la recherche était considérée comme fructueuse. Si l'utilisateur croyait à tort que l'information ne se trouvait pas dans la base de données (à l'exception des questions sans réponse), on considérait la recherche comme ayant pris fin de façon inexacte et prématurée.

L'expérimentateur a chronométré la recherche dans chacun des 16 cas. Le chronométrage débutait au moment où l'utilisateur commençait la recherche et se terminait lorsque la bonne page apparaissait sur l'écran. On pouvait également mettre fin au temps de recherche lorsque l'utilisateur indiquait à l'expérimentateur que la réponse ne se trouvait pas dans la base de données. Pour ne pas intimider les participants ou les inciter à accélérer leurs recherches, on ne leur a pas dit qu'ils seraient chronométrés. La durée moyenne des séances, y compris les instructions et les démonstrations, a été de 75 minutes.

A la fin de l'expérience on a demandé aux participants de remplir un petit questionnaire portant sur l'expérience elle-même et sur l'utilisation du Télidon en général. On a pu ainsi recueillir leurs commentaires sur l'utilisation d'un vrai système Télidon. Enfin, on a expliqué aux utilisateurs le but de l'expérience.

RÉSULTATS

Une analyse factorielle de variance à blocs réduits a démontré qu'il n'y avait pas de différence significative entre le groupe verbal et le groupe non verbal en ce qui a trait au temps de recherche ou au nombre total d'erreurs par participant, soit F (1,22) moins que 1 et F (1,22)= 1.43, p étant plus grand que .20, respectivement. Les données obtenues dans les deux groupes ont donc été condensées en un seul groupe de 24 participants. (Voir tableaux d'analyse de variance à l'annexe C.)

Si l'on tient compte du temps de recherche total, les participants ont passé en moyenne près de 160 secondes par question. Le temps de recherche moyen par type de question a été le suivant: 160 secondes pour les dix questions spécifiques, 142 secondes pour les trois questions générales et 177 secondes pour les trois questions sans réponse. Le total des temps de recherche consacrés aux questions spécifiques a pu être gonflé par l'inclusion de questions spécifiques que les participants ont abandonné avant de trouver la réponse. Pour déterminer le temps de recherche consacré aux questions spécifiques résolues, il suffit de soustraire les temps de recherche consacrés aux questions spécifiques irrésolues du total des temps de recherche consacrés aux questions spécifiques. On obtient ainsi un temps de recherche moyen de 137 secondes pour les questions spécifiques résolues et de 205 secondes pour les questions spécifiques irrésolues.

Une analyse de variance des temps de recherche moyens consacrés aux questions spécifiques résolues, aux questions générales, aux questions sans réponse et aux questions spécifiques irrésolues a révélé des écarts significatifs entre ces quatres moyennes: F (3,23) = 6.29, p moins que .001. Le test Newman-Keuls a démontré que les

questions spécifiques irrésolues demandaient beaucoup plus de temps que les questions spécifiques résolues ou les questions générales.

Le rendement des participants s'est peut-être amélioré au fur et à mesure que ceux-ci répondaient aux 16 questions. Cependant, même si le temps de recherche moyen et le nombre moyen d'erreurs commises ont diminué entre les trois premières et les trois dernières questions spécifiques, les écarts n'étaient pas significatifs.

Puisque la démarche et la conception de la présente expérience diffèrent de celles des expériences précédentes, on a établi d'autres méthodes de définition des erreurs. Une des méthodes employées pour calculer les erreurs dans la présente expérience s'appuyait sur tous les choix des participants. Toutefois, puisque la démarche suivie dans la présente expérience était différente de celle qui a été suivie au cours des expériences précédentes, on ne pouvait comparer les taux d'erreurs antérieurs au taux d'erreurs fondé sur la totalité des choix constaté dans la présente expérience. Par conséquent, afin de pouvoir comparer les résultats de la présente expérience avec ceux des expériences précédentes on a dû recourir à deux autres méthodes de définition des erreurs.

La première méthode utilisée dans la présente expérience pour calculer les erreurs est quelque peu différente des méthodes d'enregistrement des erreurs des études précédentes. Lorsqu'un participant faisait un choix "incorrect", c'est-à-dire un choix qui ne le mettait pas directement sur la voie de l'information recherchée, on comptait une erreur. Tout autre choix fait à partir d'un choix "incorrect" était également considéré comme une erreur.

Par exemple, si un utilisateur consultait le "Guide des affaires" pour obtenir un renseignement sur les "Arts", on comptait une erreur.

Chaque choix consécutif fait à l'intérieur du guide des affaires et ne menant pas à la catégorie des arts était considéré comme une erreur. De tels choix peuvent être considérés comme des erreurs puisqu'ils éloignent de plus en plus l'utilisateur de l'information recherchée. Dans une telle situation, le bon choix consiste à sélectionner la page d'index précédente en appuyant sur la touche retour arrière. De la même façon, si la page d'index visualisée comportait effectivement une rubrique correcte et que l'utilisateur choisissait de revenir à la page d'index précédente, on comptait également une erreur puisqu'il s'écartait de plus en plus de l'information recherchée.

D'après cette première définition, les participants à l'expérience ont commis 1282 erreurs au total. Le nombre d'erreurs par utilisateur a été établi à 53,42, soit une moyenne de 3,34 par problème. Les 24 utilisateurs ont consulté au total 3654 pages. Par conséquent, la probabilité d'erreur pour chaque page donnée était de 0,351. La probabilité d'un bon choix s'établit à peine à 0,65 par page d'index présentée.

Une deuxième méthode a été utilisée pour définir les erreurs. On peut alléguer que dans la première méthode certains choix des utilisateurs ne sont pas vraiment des erreurs. Par exemple, selon la première méthode, tout choix qui écarte l'utilisateur de l'information recherchée est considéré comme une erreur. Dès que l'utilisateur se trouve sur une voie qui ne mène pas directement à l'information désirée, on compte une erreur pour tout choix de rubrique présentée sur la page de menu (le choix correct d'après cette définition est de revenir en arrière). La deuxième méthode diffère de la première du fait que, pour qu'on puisse compter une erreur, il faut que chaque page de menu comporte une bonne rubrique que l'utilisateur n'a pas choisie. D'après cette méthode, il n'y a pas de choix correct ou incorrect lorsque l'utilisateur se trouve sur la mauvaise voie. Cette méthode est analogue à celle qu'ont employée Whalen et Latrémouille en 1980.

Pour calculer les erreurs selon la deuxième méthode, on a déterminé le chemin le plus court correspondant à chaque problème spécifique, c'est-à-dire le nombre minimum de pages requis pour trouver la solution. Toutes les pages qui ne se trouvaient pas sur le chemin le plus court étaient tout simplement ignorées. Si l'on considère seulement les pages qui se trouvaient sur le chemin le plus court, chaque participant pouvait faire ou bien un choix correct ou bien une erreur. Un choix était correct lorsque le participant demandait la page consécutive suivante sur le chemin le plus court le menant à l'information désirée (une telle page se trouvait nécessairement au niveau suivant de l'arbre). Tout autre choix était considéré comme une erreur. Par exemple, on comptait une erreur lorsqu'un participant choisissait une page incompatible avec le chemin le plus court ou qu'il choisissait une page du chemin le plus court, mais plus loin de la solution que la page visualisée (en appuyant sur la touche retour arrière).

D'après cette deuxième définition, les participants ont fait 443 erreurs au total. Le nombre moyen d'erreurs par utilisateur a été établi à 18,46 soit une moyenne de 1,85 erreur par question spécifique par utilisateur. Les 24 participants ont consulté au total 1312 pages sur le chemin le plus court. Par conséquent, la probabilité d'erreur par page présentée (sur le chemin le plus court) était de 0,338.

Une troisième méthode de définition des erreurs a été employée dans une expérience précédente menée par Lee et Latrémouille (1979). Cette méthode était une variante de la deuxième méthode que nous venons de décrire. Dans l'expérience menée par Lee et Latrémouille, on indiquait aux utilisateurs si leurs choix étaient exacts ou inexacts. Par conséquent, dès qu'ils avaient fait un bon choix dans une page donnée, ils ne pouvaient pas revenir au niveau précédent. Ils ne pouvaient pas non plus se tromper plus d'une fois sur le choix d'une même page. En analysant les résultats de la présente étude selon cette méthode, on n'a pas tenu compte des erreurs qui

consistaient à revenir sur une page d'index précédente après un bon choix. De plus, on a également éliminé les erreurs qui consistaient à revenir plus d'une fois sur la mauvaise page. Cette troisième méthode de calcul des erreurs s'apparente aux méthodes employées par Lee et Latrémouille et permet de comparer les taux d'erreurs des deux expériences. La présente méthode n'est toutefois pas identique à celle de l'expérience précédente. Par exemple, dans la présente expérience, les sujets, contrairement à ceux de l'expérience précédente, n'étaient pas immédiatement informés si leurs choix étaient exacts ou inexacts.

D'après cette troisième définition, les participants ont commis 275 erreurs. Le nombre moyen d'erreurs par participant a été établi à 11,46 soit une moyenne de 1,15 erreur par question spécifique par utilisateur. Les 24 participants ont consulté au total 1063 pages sur le chemin le plus court. La probabilité d'erreur par page présentée (sur le chemin le plus court) s'établit donc à 0,258 par utilisateur.

Le rapport entre le nombre de pages effectivement consultées et le minimum de pages nécessaire au repérage de l'information relative aux 10 questions spécifiques s'établit à 2.00:1. Les participants ont donc consulté deux fois plus de pages que le minimum requis pour trouver l'information. Le nombre moyen de pages consultées par type de question a été établi à 9,53 pour les questions spécifiques, 7,82 pour les questions générales et 11,58 pour les questions sans réponse. Le nombre minimum de pages requis pour répondre aux questions spécifiques a été établi en moyenne à 4,7.

Les participants ont convenablement mené à terme 72% de leurs recherches, exception faite des questions sans réponse. Inversement, 28% des recherches se sont terminées de façon inexacte et prématurée. Aucune question n'a été résolue de façon parfaite par tous les

participants, c'est-à-dire sans erreurs et sans dépassement du nombre minimum de pages requis.

Le plus haut pourcentage d'erreurs par page d'index a été vérifié dans la première page de l'arbre, soit 22% des 1282 erreurs commises. Plus de 47% des erreurs se sont produites aux deux premiers niveaux de l'arbre. Cette mesure de la qualité d'une page donnée est tout particulièrement utile pour déterminer les secteurs de la structure arborescente qu'il faut améliorer. Puisque le pourcentage d'erreurs le plus élevé correspond à la première page d'index, on pourrait améliorer grandement le rendement général des utilisateurs en améliorant cette première page. Toutefois, les utilisateurs consultent la première page beaucoup plus souvent que toute autre page de l'arbre, car, pour avoir accès à toute information dans la base de données, ils doivent nécessairement passer par la première page. Il y a donc beaucoup plus de risques de commettre une erreur dans la première page que dans toute autre qui se trouve à un niveau inférieur de l'arbre et que l'on consulte moins fréquemment. Par conséquent, une telle mesure ne reflète pas la difficulté d'une page par rapport à une autre. Il vaut mieux déterminer le pourcentage de toutes les présentations d'une page donnée qui ont entraîné des erreurs. Ainsi, la première page a été consultée 730 fois par les utilisateurs et a fait l'objet d'un choix dans chaque cas. Sur 730 choix, les utilisateurs ont commis 278 erreurs, soit un taux d'erreur de 38%. Dans une page du deuxième niveau de l'arbre, qui a été présentée 67 fois, les utilisateurs ont commis un total de 47 erreurs, soit un taux d'erreur de 70%. On en conclut que cette page particulière du deuxième niveau est pire que la page du premier niveau. Cette mesure sert à déterminer les pages d'index les mieux conçues et les pages les plus mal conçues. (Voir la liste complète des statistiques à l'annexe D).

DISCUSSION

Les résultats de la présente expérience ressemblent à beaucoup d'égards à ceux des expériences menées par Lee et Latrémouille (1979), Whalen et Latrémouille (1980) et Van Ness et Tromp (1979). Par exemple, on a remarqué dans chaque expérience que les utilisateurs éprouvaient de la difficulté à repérer l'information demandée. Les erreurs se sont produites principalement aux deux premiers niveaux de la structure arborescente (Lee et Latrémouille, 1979; Whalen et Latrémouille, 1980). L'expérience de Van Ness et Tromp a établi un rapport de deux pour un entre le nombre de pages effectivement consultées et le minimum de pages requis ainsi qu'un temps de recherche comparable de l'ordre de deux à trois minutes.

Quant à savoir si "le Télidon est facile à utiliser", les données recueillies montrent que la version originale de la base de données de démonstration du Télidon a causé certains problèmes aux utilisateurs. Le degré de difficulté éprouvé par les utilisateurs a été mesuré selon le nombre de pages consultées, le pourcentage de fins inexactes et prématurées, le nombre de problèmes résolus sans erreurs et le taux d'erreur. Le fait que les gens aient consulté deux fois plus de pages que le nombre minimum requis pour trouver les renseignements désirés peut être considéré comme un manque général d'efficacité dans le processus de recherche. En moyenne, les utilisateurs ont mis fin prématurément à leur recherche dans 28% des cas où la réponse se trouvait dans la base de données. Aucune question n'a été résolue sans erreur par tous les utilisateurs. Les probabilités d'erreurs ont été établies à 0,35, 0,34 et 0,26 d'après les trois méthodes de calcul utilisées. Les taux d'erreurs sont supérieurs à ceux établis lors des expériences menées par Lee et Latrémouille et par Whalen et Latrémouille et indiquent le degré de difficulté éprouvé par l'utilisateur dans ses recherches. Malgré la similitude des

paramètres de comparaison utilisés, il existe des écarts entre les taux d'erreurs de la présente expérience et ceux des expériences précédentes. On peut invoquer plusieurs raisons pour expliquer ces écarts. Une d'entre elles serait que la présente expérience, contrairement aux autres, a été menée sur un vrai système Télidon. On pourrait également attribuer ces écarts au fait que l'on n'a pas utilisé les mêmes bases de données ni les mêmes questions de recherche dans chaque expérience. Une dernière raison serait que, lors de l'étude menée par Lee et Latrémouille, on indiquait aux utilisateurs si leurs choix étaient corrects ou non.

Presque la moitié de toutes les erreurs se sont produites aux deux premiers niveaux de l'arbre, ce qui ressemble aux résultats des études menées par Lee et Latrémouille et Whalen et Latrémouille. Puisque les pages des deux premiers niveaux doivent être consultées beaucoup plus souvent que les pages des niveaux inférieurs (étant donné que dans toute recherche on doit nécessairement passer par la première page et ainsi de suite), on doit s'attendre à y trouver un plus grand nombre d'erreurs.

Les noms de catégorie utilisés pour désigner les différentes rubriques ont peut-être dérouté quelque peu les utilisateurs. Le questionnaire auquel les participants ont répondu après l'expérience révèle que ceux-ci, en général, estiment les rubriques trop vastes ou ambiguës (par exemple "Guide d'intérêt général" ou "Vivre au Canada").

Les commentaires recueillis au cours de l'étude de Lee et Latrémouille ont révélé que certains titres de la première page étaient trop généraux et qu'il y avait au deuxième niveau un certain recoupement des catégories et certains termes ambigus.

Les problèmes reliés à la base de données mise à l'essai dans la présente expérience montraient bien qu'il fallait faire un effort afin

de l'améliorer. Une façon d'améliorer le système de démonstration du Télidon était d'enlever les rubriques ambiguës et de les remplacer par d'autres plus claires. Les rubriques ambiguës que l'on trouve au premier et au deuxième niveaux risquent de fausser les taux d'erreurs en mettant les utilisateurs sur la mauvaise voie. Puisque la plupart des erreurs se sont produites aux deux premiers niveaux de l'arbre, c'est là qu'on a fait le plus de changements. Une expérience menée récemment par Latrémouille et Lee (1981) a démontré que l'on pouvait améliorer considérablement la première page à l'aide de tests empiriques portant sur plusieurs "page un" possibles. Cette même expérience a également révélé qu'en utilisant des descripteurs, c'est-à-dire de brèves descriptions du contenu de chaque rubrique, on améliorait grandement le rendement de l'utilisateur et aussi que ce dernier préférait de beaucoup les rubriques avec descripteurs.

Une analyse empirique des différentes méthodes de recherche pourrait également apporter de nouvelles améliorations. Parmi les différents systèmes de recherche on retrouve les systèmes à mots-clés, les systèmes hybrides à structure arborescente et mots-clés, les systèmes à structure arborescente avec renvois et les systèmes hybrides à structure ramifiée et répertoire. Ces systèmes feront l'objet de rapports ultérieurs.

Les données du présent rapport montrent que l'expérience acquise en cours d'utilisation du Télidon a suscité une diminution du temps de recherche et du nombre d'erreurs commises entre les trois premières et les trois dernières questions spécifiques, même si ces données ne sont pas significatives. Il est possible que des utilisateurs plus expérimentés auraient eu moins de difficultés à résoudre les problèmes, mais on ne peut l'affirmer avec certitude tant qu'on n'aura pas fait l'expérience.

S'il existe vraiment une différence entre le rendement de l'utilisateur "non initié" et celui de l'utilisateur "expérimenté", on

devra peut-être adopter un autre mode d'accès à l'information. Si la méthode de recherche est trop difficile, ceux qui utilisent le système pour la première fois risquent de s'en lasser très vite. D'autre part, si le système de recherche est trop simplifié, beaucoup d'utilisateurs expérimentés trouveront qu'il ne répond pas vraiment à leurs besoins. Il faudra faire un compromis entre ces deux extrêmes ou alors établir une méthode pour les non initiés et une autre pour les utilisateurs expérimentés. Par exemple, un système de recherche hybride à structure arborescente et mots-clés pourrait convenir à ces deux types d'utilisateurs en offrant une structure arborescente aux utilisateurs non initiés et un système de recherche par mots-clés aux utilisateurs expérimentés. Des systèmes de ce genre sont actuellement à l'étude.

Il est évidemment nécessaire de faire d'autres recherches dans ce domaine. Par exemple, est-ce que les gens tolèreront une probabilité d'erreurs de l'ordre de 35% ? Quel est le seuil de tolérance des utilisateurs? De plus, il faut comparer le rendement que donne l'utilisateur sur un système de recherche à structure arborescente avec son rendement sur les autres systèmes de recherche. Pour répondre à ces questions il faut absolument faire des tests empiriques. Comme nous l'avons déjà mentionné, ces systèmes seront largement répandus dans un proche avenir. Il faut donc concentrer nos efforts sur les tests empiriques des principales variables en cause. Il faudra en outre se pencher davantage sur les méthodes de recherche afin d'optimiser le procédé de consultation. Par exemple, on devrait explorer systématiquement les effets qu'entraîne l'addition de renvois et de descripteurs à l'index ramifié.

DOCUMENTATION

Latrémouille, S., et Lee, E. La conception d'index ramifiés pour le vidéotex. Recherche en comportement sur Télidon 2. Ministère des Communications, Ottawa, mai 1981.

Lee, E., et Latrémouille, S., Évaluation de la structure arborescente de l'information dans la base de données du Télidon. (Note technique N° BRG 79-12). Ministère des Communications, Ottawa, décembre, 1979. Se trouve également dans Recherche en comportement sur Télidon 1. Ministère des Communications, Ottawa, 1980, pp.228 à 236.

Van Ness, F.L., & Tromp, J.H., Is viewdata easy to use? Institute for Perception Research. Eindhoven, Netherlands: IPO Annual Progress Report, N° 14, 1979.

Whalen, T., & Latrémouille, S., L'efficacité d'un index ramifié lorsque l'existence de données est incertaine. (Note technique N° BRIC 80-3). Ministère des Communications, Ottawa. Septembre 1980.

ANNEXE A

INSTRUCTIONS

Au cours de la présente expérience vous ferez connaissance avec le Télidon. Le Télidon est un système de communication visuelle interactif qui permet au public d'avoir accès à des sources d'informations stockées en ordinateur. L'utilisateur sélectionne les pages d'information contenues dans la base de données à l'aide d'un clavier. La page choisie apparaît ensuite sur l'écran d'un téléviseur.

La base de données Télidon contient actuellement une foule de renseignements. On y trouve entre autres, des listes de restaurants d'Ottawa, des nouvelles du monde des affaires et des inscriptions immobilières. La base de données est agencée en une structure arborescente qui ressemble à un arbre inversé. Cette structure arborescente se ramifie à partir d'une seule racine pour former un ensemble de niveaux hiérarchiques descendants, chacun ayant des points nodaux qui le relient à un autre niveau.

Il existe deux types de pages d'affichage: les pages d'index et les pages documentaires. Les pages d'index divisent une catégorie en plusieurs sous-catégories pour indiquer à l'utilisateur, qui fait un choix à chaque niveau de l'arbre, où chercher la page suivante appropriée. Les pages documentaires contiennent l'information qui intéresse directement l'utilisateur. En choisissant une rubrique donnée dans une page d'index située à un point nodal, l'utilisateur passe à un point nodal du niveau inférieur suivant. Au fur et à mesure qu'on descend dans l'arbre, l'information devient de plus en plus précise jusqu'à ce que le document désiré soit repéré.

Votre travail consistera à repérer dans la base de données les documents désirés. On vous posera 16 questions qui détermineront l'information qu'il faut repérer.

La plupart des réponses se trouvent dans la base de données. Dans certains cas, la réponse peut ne pas s'y trouver. Votre recherche prendra fin au moment où vous repérerez l'information désirée ou lorsque vous croirez que l'information ne se trouve pas dans la base de données.

Au cours de vos recherches dans la base de données vous remarquerez que certaines pages d'index comportent une ou plusieurs rubriques fermées. Cela signifie que vous ne pouvez pas avoir accès à l'information de ces rubriques puisqu'elles ne contiennent actuellement aucune information. Aux fins de la présente expérience, les réponses à certaines questions peuvent se trouver sous une rubrique fermée. Si c'est le cas et que vous croyez trouver la réponse sous cette rubrique, vous pouvez alors "choisir" la rubrique en question en disant son nom ou son numéro. Si vous choisissez une rubrique fermée qui ne contient pas la bonne réponse, on vous le fera savoir et vous pourrez continuer votre recherche.

Pendant que vous cherchez l'information désirée nous aimerions vous entendre décrire votre façon de procéder. Nous voulons savoir à quoi vous pensez pendant que vous faites vos recherches. Par exemple, nous aimerions savoir pourquoi vous choisissez une rubrique plutôt qu'une autre ou encore pourquoi vous avez choisi de chercher dans une partie particulière de l'arbre. En d'autres termes, il s'agit de penser tout haut.

Nous aimerions également que vous donniez les raisons ou les stratégies qui déterminent vos choix avant d'appuyer sur les touches du clavier. Nous vous prions de motiver tous vos choix même si cela

vous semble superflu ou inutile. Toutes les raisons sont importantes pour nous. Vos commentaires seront enregistrés au cours de cette expérience.

Vous ferez probablement quelques erreurs au cours de vos recherches. Les erreurs n'entraînent pas de pénalités et elles seront plutôt attribuables à la conception du système qu'à vous-mêmes. Soyez donc bien à l'aise et ne vous inquiétez pas si jamais vous faites une erreur.

Avant de commencer l'expérience nous vous montrerons comment utiliser le clavier du Télidon et vous donnerons un exemple de ce que vous avez à faire. Si vous avez des questions, ne vous gênez pas pour les poser.

ANNEXE B

Questions de recherche

1. Quelles sont les principales divisions territoriales de la Suisse?

2. Vous voulez faire réparer votre voiture. Trouvez des renseignements à ce sujet.

3. Où pouvez-vous trouver des renseignements sur l'achat d'une propriété à l'extérieur d'Ottawa?

4. Combien de stations radio y a-t-il à Ottawa?

5. Vous voulez acheter un réfrigérateur. Que devez-vous savoir avant de faire votre achat?

6. Quelqu'un vous a parlé d'un restaurant italien appelé Calabria. Quelle est l'adresse de ce restaurant?

7. Vous venez de déménager à Ottawa et vous voulez savoir à quelle école envoyer vos enfants cet automne.

8. Trouvez des renseignements sur la géographie de Terre-Neuve.

9. Quelles sont les limites de la circonscription électorale provinciale de Riel à Winnipeg?

10. Que vous prédit votre horoscope pour aujourd'hui?

11. Quels renseignements pouvez-vous obtenir sur les arts?

12. Vous venez de gagner à la loterie Wintario et vous voulez savoir comment investir une partie de vos gains.

13. Que se passe-t-il dans le monde des sports?

14. Comment jouer au jeu trigonométrique?

15. Quel portefeuille détient Robert Smith dans le cabinet ministériel de la Saskatchewan?

16. Vous voulez consulter une liste des voitures 1980 à vendre.

ANNEXE C

TABLEAUX DES ANALYSES DE LA VARIANCE

TABLEAU DE L'ANALYSE DE LA VARIANCE A BLOCS RÉDUITS:
TEMPS DE RECHERCHE

Source	SS (Sommes de carrés)	df (degrés de liberté)	MS (carrés moyens)	F	SIGNI-FICATI-VITÉ
1. INTERGROUPES	1060755,13	23	48459,13		
2. VERBALISATION	255,13	1	255,13	0,01	0,943
3. ERREUR	1060500,00	22	48204,00		
4. INTRAGROUPES	5912310,00	360	94851,00		
5. QUESTION	967460,00	15	64498,00	4,52	0,001
6. VERBALISATION x QUESTION	233650,00	15	15577,00	1,09	0,363
7. ERREUR	4711200,00	330	14276,00		
8. TOTAL	6973000,00	383			

VARIABLE: NOMBRE D'ERREURS

Source	SS (Sommes de carrés)	df (degrés de liberté)	MS (carrés moyens)	F	SIGNI-FICATI-VITÉ
1. INTERGROUPES	1062,11	23	110,30		
2. VERBALISATION	65,01	1	65,01	1,43	0,224
3. ERREUR	997,10	22	42,32		
4. INTRAGROUPES	9137,10	360	194,58		
5. QUESTION	2267,70	15	151,18	7,66	0,001
6. VERBALISATION x QUESTION	354,91	15	23,66	1,20	0,271
7. ERREUR	6514,70	330	19,74		
8. TOTAL	10199,00	383			

TABLEAU D'ANALYSE DE LA VARIANCE PAR BLOCS AVEC RÉPARTITION AU HASARD
VARIABLE: TEMPS DE RECHERCHE

Source	SS (Sommes de carrés)	df (degrés de liberté)	MS (carrés moyens)	F	SIGNIFICATIVITÉ
1. TRAITEMENTS[1] DIFFÉRENTS	98047,00	3	32682,00	6,29	0,001
2. INTERGROUPES	408710,00	23	17770,00		
3. ERREUR	358320,00	69	5193,10		
4. TOTAL	865080,00	95			

[1] Type de question: Spécifique résolue, spécifique irrésolue, générale, sans réponse.

ANNEXE D

RÉSULTATS DE L'EXPÉRIENCE DE RECHERCHE

TEMPS DE RECHERCHE MOYEN
n=16 QUESTIONS = 159,99 s
n=10 QUESTIONS (SPÉCIFIQUES) = 160,42 s
n=3 QUESTIONS (GÉNÉRALES) = 141,77 s
n=3 QUESTIONS (SANS RÉPONSE) = 176,81 s

SP (résolu) = 137,30 s
SP (irrésolu)= 204,93 s

PREMIÈRE MÉTHODE DE DÉTERMINATION DU NOMBRE D'ERREURS
NOMBRE TOTAL D'ERREURS = 1282
NOMBRE MOYEN D'ERREURS PAR UTILISATEUR = 1282/24 = 53,42
NOMBRE MOYEN D'ERREURS PAR UTILISATEUR PAR QUESTION = 3,34
NOMBRE TOTAL DE PAGES CONSULTÉES = 3654
PROBABILITÉ D'ERREUR = 1282/3654 = 0,35 (0,3508)

DEUXIÈME MÉTHODE DE DÉTERMINATION DU NOMBRE D'ERREURS
NOMBRE TOTAL D'ERREURS = 443
NOMBRE MOYEN D'ERREURS PAR UTILISATEUR = 443/24 = 18,46
NOMBRE MOYEN D'ERREURS PAR UTILISATEUR PAR QUESTION SPÉCIFIQUE = 1,85
NOMBRE TOTAL DE PAGES CONSULTÉES = 1321
PROBABILITÉ D'ERREUR = 443/1321 = 0,34 (0,3377)

TROISIÈME MÉTHODE DE DÉTERMINATION DU NOMBRE D'ERREURS
NOMBRE TOTAL D'ERREURS = 275
NOMBRE MOYEN D'ERREURS PAR UTILISATEUR =275/24 = 11,46
NOMBRE MOYEN D'ERREURS PAR UTILISATEUR PAR QUESTION SPÉCIFIQUE = 1,15
NOMBRE TOTAL DE PAGES CONSULTÉES = 1063
PROBABILITÉ D'ERREUR = 275/1063 = 0,26 (0,2587)

TOTAL DES CHOIX PORTANT SUR DES QUESTIONS SPÉCIFIQUES = 2277
TOTAL DES QUESTIONS SPÉCIFIQUES = 239
2277/239 = 9,53 PAGES PAR QUESTION SPÉCIFIQUE
TOTAL DES CHOIX PORTANT SUR DES QUESTIONS GÉNÉRALES = 555
TOTAL DES QUESTIONS GÉNÉRALES = 71
575/71 = 7,82 PAGES PAR QUESTION GÉNÉRALE
TOTAL DES CHOIX PORTANT SUR DES QUESTIONS SANS RÉPONSE = 822
TOTAL DES QUESTIONS SANS RÉPONSE = 71
822/71 = 11,58 PAGES PAR QUESTION SANS RÉPONSE

RAPPORT ENTRE LE NOMBRE DE PAGES EFFECTIVEMENT CONSULTÉES ET LE NOMBRE MINIMUM DE PAGES CONSULTÉES (n = 10 QUESTIONS SPÉCIFIQUES)
TOTAL DES CHOIX PORTANT SUR DES QUESTIONS SPÉCIFIQUES = 2277
NOMBRE MINIMUM TOTAL DE PAGES CONSULTÉES - QUESTIONS SPÉCIFIQUES = 1138
2277/1138 = 2,00

NOMBRE MOYEN DE RECHERCHES FRUCTUEUSES (n = 13 QUESTIONS SANS RÉPONSE EXCLUES)
NOMBRE TOTAL DE RECHERCHES FRUCTUEUSES = 222
N = 24
224/24 = 9,25 RECHERCHES FRUCTUEUSES PAR SUJET

POURCENTAGE DE RECHERCHES FRUCTUEUSES (n = 13 QUESTIONS SANS RÉPONSE EXCLUES)
NOMBRE TOTAL DE RECHERCHES FRUCTUEUSES = 222
TOTAL DES QUESTIONS SPÉCIFIQUES ET DES QUESTIONS GÉNÉRALES = 310
222/310 = 72% (0,7161)

NOMBRE MOYEN DE RECHERCHES TERMINÉES PRÉMATURÉMENT (n = 13 QUESTIONS SANS RÉPONSE EXCLUES)
NOMBRE TOTAL DE FINS PRÉMATURÉES ET INCORRECTES = 88
N = 24
88/24 = 3,67 FINS PRÉMATURÉES ET INCORRECTES PAR SUJET

POURCENTAGE DES RECHERCHES TERMINÉES PRÉMATURÉMENT (n = 13 QUESTIONS SANS RÉPONSE EXCLUES)
NOMBRE TOTAL DE FINS PRÉMATURÉES INCORRECTES = 88
TOTAL DES QUESTIONS SPÉCIFIQUES ET DES QUESTIONS GÉNÉRALES = 310
88/310 = 28% (0,2839)

NOMBRE MOYEN D'ERREURS PAR QUESTION

QUESTION	MOYENNE SELON LA MÉTHODE 1	MOYENNE SELON LA MÉTHODE 2	MOYENNE SELON LA MÉTHODE 3
1	1,417	0,500	0,458
2	4,292	1,833	0,875
3	7,083	1,958	1,333
4	2,458	1,292	0,583
5	9,917	2,583	1,667
6	3,917	2,917	1,917
7	1,250	0,792	0,750
8	4,458	3,167	1,417
9	5,052	2,333	1,625
10	2,875	1,083	0,958
11	0,500		
12	1,042		
13	1,625		
14	1,708		
15	4,542		
16	1,792		

NOMBRE DE QUESTIONS RÉSOLUES CORRECTEMENT PAR TOUS LES SUJETS = 0,00
NOMBRE DE QUESTIONS RÉSOLUES CORRECTEMENT = 63
POURCENTAGE DES QUESTIONS RÉSOLUES CORRECTEMENT = 63/239 = 26%
(0,2636)
POURCENTAGE DES QUESTIONS FAISANT L'OBJET D'AU MOINS UNE ERREUR = 59%
(0,5936)

POURCENTAGE D'ERREURS SUR UNE PAGE DONNÉE:

NIVEAU	PAGE	% D'ERREUR	
1	PAGE-TITRE	22	278/1282 = 0,2168
2	1,0-9,0	26	330/1282 = 0,2574
2	1,0	13	164/1282 = 0,1279
2	2,0	07	86/1282 = 0,0671
2	3,0	04	47/1282 = 0,0367
2	4,0	01	10/1282 = 0,0078
2	5,0	0,5	7/1282 = 0,0055
2	6,0	0,2	3/1282 = 0,0023
2	7,0	0,3	4/1282 = 0,0031
2	8,0	0,2	2/1282 = 0,0016
2	9,0	0,5	7/1282 = 0,0055

NIVEAU 1 + 2 (PAGE-TITRE + 1,0-9,0) = 0,47 608/1282 = 0,4743

POURCENTAGE DE CHOIX ERRONÉS FAITS SUR UNE PAGE DONNÉE

NIVEAU	PAGE	% D'ERREUR	
1	PAGE-TITRE	38	278/730 = 0,3808
2	1,0-9,0	38	330/879 = 0,3754
2	1,0	34	146/476 = 0,3445
2	2,0	61	86/142 = 0,6056
2	3,0	70	47/67 = 0,7015
2	4,0	33	10/30 = 0,3333
2	5,0	09	7/79 = 0,0886
2	6,0	30	3/10 = 0,3000
2	7,0	19	4/21 = 0,1905
2	8,0	05	2/39 = 0,0513
2	9,0	47	7/15 = 0,4667

NIVEAU 1 + 2 (PAGE-TITRE + 1,0-9,0) = 0,38 608/1609 =0,3779

CHAPITRE 4

LA CONCEPTION D'INDEX RAMIFIÉS DU VIDÉOTEX: UTILISATION DE DESCRIPTEURS ET AMÉLIORATION DES INDEX PAGE PAR PAGE

Susane Latrémouille et Eric Lee

RÉSUMÉ

Actuellement, dans la plupart des systèmes vidéotex, la recherche documentaire fait appel à une structure ramifiée ou à la sélection de menus sous une forme ou une autre. La théorie de l'information est muette sur la façon de construire des index ramifiés. Les fournisseurs d'informations qui désirent établir des index ramifiés doivent s'en remettre principalement à leur intuition. La présente étude vise à faire des recommandations et à proposer des lignes directrices à cet égard. Les deux expériences dont on fait état dans le présent document ont pour objet: a) de montrer l'application d'une autre méthode de modification d'index ramifiés, c'est-à-dire la méthode page par page; b) d'évaluer l'effet de l'addition de descripteurs (brèves descriptions de chaque rubrique) sur les préférences et le rendement des utilisateurs; c) de déterminer s'il y a, chez les initiés ou chez les non-initiés, des préférences communes quant aux "pages 1" possibles; d) de déterminer la relation entre le rendement et les préférences des utilisateurs quant aux pages 1; et e) de déterminer, parmi douze pages 1 possibles, celle qui est la mieux adaptée à la base de données expérimentale du Télidon du ministère des Communications.

Dans la première expérience, dix personnes initiées au vidéotex ont classé par ordre de préférence douze pages 1 (pages supérieures de l'index) en vue d'une base de données expérimentale. Six de ces douze pages avaient été produites par des experts travaillant de façon indépendante. Les autres pages comprenaient la page originale, la page modifiée résultant de modifications apportées à la page originale lors d'une expérience précédente et la page produite par un indexateur professionnel. Chacune des trois dernières pages était présentée en deux versions: une page avec descripteurs et une page sans descripteurs. Nous n'avons pour ainsi dire constaté aucune préférence commune chez les initiés, car chacun d'eux avait une opinion

différente de ce qui constituait une meilleure page. Aucune des douze pages n'a donc été jugée, de façon fiable, vraiment supérieure aux autres.

Dans la deuxième expérience, vingt personnes non initiées au vidéotex ont accompli deux tâches: la première, une recherche, et la deuxième, un classement, par ordre de préférence, des douze pages d'index ayant servi à la première expérience. Au cours de la recherche, on posait à l'utilisateur dix questions pour chacune des douze pages d'index (soit 120 questions au total). Celui-ci répondait en choisissant une rubrique dans la page d'index affichée à l'écran. Les utilisateurs non initiés ont fait preuve de cohérence tant dans le classement par ordre de préférence que dans le nombre d'erreurs de sélection commises par rapport à chaque page. Certaines pages ont été jugées vraiment supérieures à d'autres. C'est pour les pages qui contenaient des descripteurs qu'il y a eu le moins d'erreurs, ces pages ayant en outre été jugées préférables à celles sans descripteurs. L'addition de descripteurs a diminué le nombre d'erreurs de moitié.

RECOMMANDATIONS ET CONCLUSIONS

1. La mise à l'essai d'un index ramifié peut servir à faciliter de façon significative la recherche documentaire du non-initié. On recommande de mettre à l'épreuve les pages 1 d'index si les ressources dont on dispose sont rares.

2. Les descripteurs réduisent sensiblement les erreurs en cours de recherche et augmentent la préférence des utilisateurs.

3. Étant donné que les initiés ne s'entendaient pas, pour ainsi dire, sur les caractéristiques d'une page d'index bien conçue, il convient d'avoir recours à un échantillon de la population cible (dans ce cas, les non-initiés) pour tester les pages d'index.

4. Les préférences et le nombre d'erreurs des utilisateurs sont deux éléments qui permettent de mesurer la qualité des pages d'un index ramifié. Dans le cas des non-initiés, il existe une corrélation très étroite entre ces deux mesures lorsque les préférences sont mises en lumière après les résultats de l'expérience. Les non-initiés ont fait preuve d'une grande cohérence tant sur le plan des préférences que sur celui du rendement. En outre, il y a des différences notables et incontestables tant dans le nombre d'erreurs que sur le plan du rendement pour les différentes pages d'index. Des recherches plus poussées sont nécessaires pour déterminer si les préférences suffisent en tant que mesure de la qualité des pages d'index aux fins d'améliorer les index ramifiés.

5. Il convient d'éviter des noms de rubriques ambigus et trop généraux comme "guide d'intérêt général" et "divers".

6. Chaque page devrait comporter au moins quatre rubriques.

7. On doit chercher à utiliser des rubriques à peu près équivalentes quant au nombre de renseignements auxquels elles renvoient sur chaque page.

8. Il importe d'organiser les pages selon une certaine structure (par ordre alphabétique ou selon la fréquence des consultations).

INTRODUCTION

Actuellement, dans la plupart des systèmes vidéotex, la recherche documentaire fait appel à une structure ramifiée ou à une sélection de menus sous une forme ou une autre. La théorie contemporaine de l'information ne vient pas vraiment en aide au constructeur d'index ramifiés. Jusqu'à maintenant, les théoriciens de l'information ont établi des systèmes de recherche par mots repères pour des livres et des articles, mais ces systèmes ne sont pas adaptés à des renseignements d'intérêt général que l'on retrouve ailleurs que dans les livres. De plus, pareils systèmes s'adressent habituellement aux bibliothécaires et non aux profanes ou au grand public.

En raison de cette carence de la théorie de l'information, les fournisseurs d'informations et d'autres personnes ont dû s'en remettre principalement à leur intuition pour établir des index ramifiés. Des expériences récentes (Lee & Latrémouille, 1979; Whalen & Latrémouille, 1980; McEwen, 1981; Van Ness & Tromp, 1979) ont démontré que le non-initié peut éprouver des difficultés énormes lorsqu'il effectue une recherche documentaire dans ces index ramifiés de première génération. Il faut trouver des méthodes et des techniques pour améliorer ces index, afin de réaliser le potentiel des index ramifiés aux fins de la recherche documentaire au moyen du vidéotex.

La série d'expériences décrites ci-après avait pour objet d'étudier des méthodes pouvant améliorer la structure arborescente du point de vue de l'utilisateur. Le principal objectif des chercheurs était de faire des recommandations à l'intention des fournisseurs d'informations et autres personnes qui auront à construire des index ramifiés pour des bases de données. Ces recommandations devraient permettre à l'indexateur de mettre au point un système d'extraction de données plus efficace et plus satisfaisant pour l'utilisateur.

Des études précédentes sur les structures arborescentes (Lee & Latrémouille, 1979; Van Ness & Tromp, 1979; McEwen, 1981) révèlent que la plupart des erreurs (50% ou plus) commises par le non-initié lors d'une recherche dans un index ramifié portent sur les deux niveaux supérieurs de l'arbre. En fait, le nombre d'erreurs par page est le plus élevé pour le niveau supérieur et il décroît avec les niveaux, c'est-à-dire qu'on retrouve le plus grand nombre d'erreurs dans la page-titre, puis dans la deuxième page, la troisième page et ainsi de suite. Ceci est attribuable, en partie, à la fréquence de consultation des pages des niveaux supérieurs de l'index. En effet, la première page de l'index doit être consultée au début de chaque séance de recherche documentaire, tandis que les pages du deuxième niveau ne servent que pour certaines séances (plus précisément, celles au cours desquelles les renseignements demandés sont en fait classés sous l'une des rubriques d'une page du deuxième niveau). Si les ressources dont on dispose pour améliorer un index ramifié donné sont restreintes, il peut être logique de concentrer ces ressources là où elles sont le plus susceptibles d'améliorer l'arbre du point de vue de l'utilisateur. Ce sont les pages d'index comportant le grand nombre d'erreurs qui ont le plus grand potentiel d'amélioration. Il s'ensuit qu'il faut faire une utilisation décroissante des ressources, en commençant par les pages d'index dont le taux d'erreurs est le plus élevé et en terminant par celles dont le taux est le plus faible. Comme la plupart des erreurs portent sur la page d'index supérieure, c'est cette page qu'il faut améliorer en premier. La méthode préconisée dans les présentes expériences consiste donc à améliorer l'arbre page par page et non en s'attaquant à l'arbre globalement comme c'était le cas dans l'étude de Lee & Latrémouille (1979). Dans le présent rapport, cette méthode d'amélioration de la structure de l'arbre sera dénommée méthode d'amélioration page par page.

Une des principales causes des erreurs commises par les utilisateurs effectuant une recherche dans un index ramifié provient de l'ambiguïté des rubriques mêmes (Lee & Latrémouille, 1979). Souvent, les rubriques figurant sur une même page d'index se recoupent passablement quant à la nature des données que les utilisateurs comptent trouver répertoriées sous chacune. De même, certains types d'informations semblent ne figurer dans aucune des rubriques de la page d'index présentée, même si, en fait, ces informations se retrouvent dans la catégorie générale, ou page-titre, qui regroupe toutes les rubriques de cette page. Dans un cas comme dans l'autre, la nature des termes utilisés est la cause des erreurs commises par l'utilisateur parce que le nom de la catégorie (rubrique) n'indique pas toujours clairement quels renseignements sont regroupés sous celle-ci. Il serait possible de réduire cette ambiguïté en fournissant une description complète de toutes les informations qu'englobe chaque catégorie (rubrique). Cependant, l'utilisateur aurait alors à lire pratiquement toutes les inscriptions de la base de données chaque fois qu'il entreprendrait une nouvelle recherche documentaire. Donc, les descriptions complètes ne sont absolument pas pratiques, même si elles permettraient d'éviter toute ambiguïté d'interprétation des rubriques. Cependant, de brèves descriptions des articles répertoriés sous chaque rubrique pourraient peut-être faciliter de façon significative la recherche documentaire du profane puisqu'elles réduiraient sensiblement l'ambiguïté intrinsèque des rubriques. Ces descriptions peuvent prendre deux formes: une description courte et abrégée du contenu de chaque catégorie (rubrique) ou une liste des rubriques regroupées sous une étiquette de catégorie donnée, que l'on retrouve au niveau inférieur suivant de l'arbre. Nous parlerons, dans un cas comme dans l'autre, de descripteurs. Néanmoins, l'addition de descripteurs sur chaque page de menu comporte aussi des inconvénients. L'addition de descriptions, même brèves, à chaque rubrique d'une page de menu peut accroître non seulement le temps de transmission et d'affichage des données en

provenance de l'ordinateur central, mais aussi le temps de lecture de ces données supplémentaires sur chaque page de menu. Elle peut aussi se traduire par des coûts accrus d'introduction et de transmission des données. Néanmoins, ces dépenses supplémentaires n'influent pas directement sur l'utilisateur et ne seraient probablement pas très grandes comparativement aux coûts résultant de recherches infructueuses.

Les effets de descripteurs peuvent donner lieu à plusieurs prédictions différentes. Ainsi, il est possible que des descriptions plus étoffées réduisent l'ambiguïté des rubriques (et du même coup les pertes de temps occasionnées par des cheminements de recherche erronés). En outre, il pourrait y avoir augmentation concomitante du temps requis pour la lecture et l'affichage des données. Donc, dans l'intérêt de l'utilisateur, on pourrait en venir à un compromis entre, d'une part, des temps de recherche plus courts du fait que l'utilisateur commettrait moins d'erreurs et, d'autre part, des temps de recherche plus longs en raison du temps de lecture et d'affichage également plus longs. La relation entre le temps de recherche de l'utilisateur (et sa satisfaction) et l'ampleur des descriptions peut s'exprimer au moyen d'une fonction curviligne comme celle qui est présentée à la figure 1. Le graphique montre qu'un nombre moyen de descriptions devrait réduire au minimum le temps de recherche et porter au maximum la satisfaction de l'utilisateur. Un nombre trop grand ou trop faible de descriptions se traduit par des temps de recherche beaucoup plus longs et un degré de satisfaction bien moindre de l'utilisateur.

D'autre part, si l'on a recours à des descripteurs, l'utilisateur peut interrompre sa lecture des pages d'index plus tôt et faire un choix plus rapidement. Du fait qu'ils clarifient le contenu de l'information regroupée sous chaque rubrique, les descripteurs

devraient permettre à l'utilisateur de repérer plus facilement la rubrique voulue. A défaut de descripteurs, l'utilisateur devra peut-être lire toutes les rubriques présentées sur une page d'index et consacrer quand même beaucoup de temps à choisir entre des rubriques essentiellement ambiguës. Des descripteurs pourraient donc réduire tant les erreurs que le temps de recherche. Une troisième possibilité existe: les utilisateurs pourraient rejeter entièrement l'utilisation de descripteurs, préférant une structure ramifiée composée uniquement de rubriques. (Cependant, l'affichage des descripteurs pourrait suivre l'affichage des rubriques. Ainsi, les utilisateurs expérimentés devraient pouvoir lire les rubriques avant l'affichage des descripteurs.) Seules des recherches empiriques permettront de répondre à la question.

La présente étude visait principalement à déterminer l'effet de courtes descriptions sur le rendement et les préférences des utilisateurs. Elle avait aussi, comme objectif secondaire, d'évaluer l'homogénéité des préférences et du rendement des non-initiés concernant les premières pages d'index, ainsi que d'apprécier la cohérence dont les initiés avaient fait preuve dans l'établissement de leurs préférences. Si les préférences et le rendement des non-initiés avaient été divergents vis-à-vis des différents choix de pages d'index du niveau supérieur, la tentation aurait été grande de conclure qu'aucun arbre ne pouvait répondre aux besoins de la plupart des utilisateurs. Par ailleurs, si les préférences et le rendement des utilisateurs devaient concorder sensiblement, il devenait assez évident a) que certaines pages d'index sont de loin préférables à d'autres et b) qu'il est possible de concevoir des pages d'index pouvant satisfaire aux besoins de la plupart des utilisateurs.

Des expériences précédentes (Lee & Latrémouille, 1979; Whalen & Latrémouille, 1980; McEwen, 1981) indiquent que les gens font souvent le mauvais choix dans la page 1 actuelle de la base de données

expérimentale du Télidon (septembre 1980). Sur le plan pratique, les expériences visaient entre autres à trouver pour cette base de données une page 1 qui réduirait le nombre d'erreurs. Donc, un objectif pratique de ces présentes expériences consistait à trouver une nouvelle page 1 suscitant le moins d'erreurs possible. Une page 1 améliorée serait particulièrement utile pour les démonstrations du Télidon, puisque cette page est presque toujours présentée à l'auditoire.

La recherche documentaire par choix de menus fait aussi intervenir une autre question: la relation entre les préférences des utilisateurs et leur rendement. Si préférences et rendement sont étroitement liés, on peut substituer au rendement les jugements de préférence comme mesure de la qualité des pages d'index. Comme il est généralement plus facile et moins coûteux de recueillir des données sur les préférences que sur le rendement, la mise à l'épreuve et la modification des index ramifiés pourraient être passablement simplifiées. La préférence des utilisateurs est-elle strictement inversement proportionnelle au nombre d'erreurs? Mis à part le rendement, quels sont les facteurs qui influent sur la perception qu'ont les utilisateurs des pages d'index ramifié? S'il n'y a aucune relation entre les préférences et le rendement des utilisateurs, de quelle façon devrait-on modifier les structures ramifiées? En tenant compte des préférences ou du rendement? C'est pour répondre à ces questions que la présente étude a été faite.

En résumé, les objectifs de la présente étude sont: a) de démontrer l'application d'une méthode de rechange pour modifier les index ramifiés, c'est-à-dire la méthode d'amélioration page par page; b) d'évaluer l'effet de l'addition de descripteurs (brèves descriptions de chaque rubrique) sur les préférences et le rendement des utilisateurs; c) de déterminer le degré d'uniformité des

préférences des initiés et des non-initiés concernant les pages 1 d'index possibles; d) de déterminer la relation qui existe entre le rendement des utilisateurs et leurs préférences concernant les pages 1 d'index; et e) de déterminer, entre douze pages 1 possibles, celle qui convient le mieux à la base de données expérimentale du Télidon du ministère des Communications.

Deux expériences ont été conçues et menées pour répondre aux objectifs décrits ci-dessus. A partir de ces recherches empiriques, nous avons formulé des recommandations à l'intention des fournisseurs d'informations et d'autres personnes qui ont à concevoir et à construire des index ramifiés.

EXPÉRIENCE I

Comme la plupart des erreurs commises portaient sur la page du niveau supérieur (page 1) de l'index ramifié (Lee & Latrémouille, 1979, McEwen, 1981), l'expérience visait à trouver une page 1 convenant mieux à la base de données expérimentale du Télidon. En supposant que les résultats escomptés soient atteints, la modification de la première page d'index devait résulter, globalement, en une amélioration considérable du rendement et de la satisfaction des utilisateurs ayant recours à la méthode de recherche ramifiée.

Douze versions différentes de la première page d'index ont été produites. Nous avons pris comme point de départ la page 1 originale parce que toute page qu'on y substituerait devait présenter de meilleures caractéristiques et parce que nous possédions déjà des données sur le rendement des utilisateurs par rapport à cette page (Lee & Latrémouille, 1979). Nous avons aussi conservé la première version modifiée de la page originale, pour les mêmes raisons. Pour obtenir cette première version modifiée, nous avons identifié les cas où des non-initiés, représentatifs du grand public, choisissaient des rubriques inappropriées dans l'arbre original, puis nous avons demandé à un autre groupe de sujets semblables de reclasser et de renommer les rubriques qui avaient induit le premier groupe en erreur (Lee & Latrémouille, 1981). Nous avons choisi des profanes plutôt que des spécialistes parce que nous croyions que les mots et classements qu'ils utiliseraient se rapprocheraient davantage de ceux qui viendraient le plus naturellement à l'esprit du non-spécialiste. En outre, un arbre fondé entre autres sur les classements établis par un échantillon d'individus serait plus représentatif qu'un autre fondé sur les classements établis par une seule personne. Nous avons supposé que les utilisateurs du Télidon feraient moins de mauvais choix si la terminologie et les catégories employées pour l'établissement de la structure choix si la terminologie et les

catégories employées pour l'établissement de la structure arborescente correspondaient le plus possible à celles qui leur viendraient le plus facilement à l'idée.

La troisième version de la page originale a été réalisée par un indexateur professionnel (bibliothécaire). Trois autres versions de la page originale ont été obtenues en ajoutant des descripteurs à chacune des rubriques de la page originale, de la page modifiée et de la page réalisée par l'indexateur professionnel. Ces six versions de la page 1 constituaient les éléments d'un modèle intégré au modèle d'ensemble à douze pages. Le modèle intégré comportait deux facteurs: le type de page 1 (pages originale, modifiée ou réalisée par l'indexateur professionnel) et les descripteurs (présence ou absence de descripteurs). Ce modèle intégré a été spécialement construit pour vérifier l'utilité d'ajouter des descripteurs aux pages d'index. Les six autres pages 1 ont été produites par des experts. La conception et l'organisation variaient considérablement d'une page à l'autre.

La première expérience présentée dans cette étude avait pour objet d'examiner les préférences des spécialistes du Télidon (des personnes connaissant bien le Télidon) quant aux douze versions de la page 1 décrites précédemment. On a demandé à chaque spécialiste de classer qualitativement ces douze versions.

MÉTHODE

SUJETS. Nous avons demandé à dix personnes connaissant assez bien la base de données du Télidon de participer à l'expérience. L'échantillon comprenait cinq anglophones et cinq personnes bilingues (parlant le français et l'anglais). Tous travaillaient régulièrement au projet Télidon et peuvent de ce fait être considérés comme des spécialistes en matière d'utilisation du vidéotex.

STIMULI. Les stimuli consistaient en douze versions de la page 1 (voir Annexe A), chacune étant dactylographiée sur une fiche blanche (de 12,5 X 20,5 cm). Chaque page devait servir comme première page d'un index ramifié utilisé pour la recherche documentaire dans une même base de données expérimentale. Aux fins de comparaison, nous avons pris la page d'index du niveau supérieur de l'arbre original et la page correspondante, modifiée par expérimentation (Lee & Latrémouille, 1979). Un bibliothécaire professionnel travaillant dans le domaine de l'indexation a réalisé une troisième version de la page 1. Aucune de ces trois pages ne contenait de descripteurs (brèves descriptions du contenu des rubriques). Un ensemble correspondant à ces trois pages, mais contenant chacune des descripteurs a été produit. Six autres pages 1 ont été constituées par des experts travaillant de façon indépendante. Ces experts n'étaient pas des indexateurs professionnels, mais des personnes connaissant bien les structures ramifiées existantes du vidéotex. Ils comprenaient aussi bien des spécialistes du comportement humain que des spécialistes des bases de données informatiques. Les experts qui ont réalisé les pages d'index dénommées expert 1:ND, expert 4:ID et expert 5:ND ont eu recours à leur propre intuition, aux observations d'utilisateurs et aux résultats d'études précédentes sur les structures ramifiées. Dans le cas de la page expert 2:ND, le constructeur a remanié une page 1 précédente dans laquelle les

utilisateurs avaient commis de nombreuses erreurs. Il a donc modifié systématiquement le classement et les titres des rubriques qui avaient induit les utilisateurs en erreur. Une méthode passablement différente a été employée pour constituer la page expert 6:ND. Au lieu de présenter des rubriques qui décrivaient le contenu d'une catégorie d'informations, cette page affichait des rubriques axées sur des actions ou sur le temps.

DÉMARCHE. Nous avons disposé les douze versions de la page 1 au hasard avant de les présenter à chaque personne. Nous avons demandé à chacune de classer ces pages en ordre de préférence pour nous permettre de déterminer celles qui convenaient à une recherche documentaire au moyen du Télidon et celles qui n'y étaient pas adaptées. Pour faciliter l'évaluation, nous avons demandé aux participants de séparer les douze pages en deux groupes: les pages acceptables et les pages de qualité médiocre. Les participants ont ensuite classé qualitativement les pages de chaque groupe, de la meilleure à la moins bonne.

RÉSULTATS ET DISCUSSION

Le tableau 1 montre les rangs moyens et normalisés de chaque page 1 évaluée selon les préférences des utilisateurs. (Les rangs accordés par chaque personne figurent à l'Annexe C.). Il n'y a pas d'uniformité quant aux préférences des sujets, le coefficient de concordance de Kendall, $\underline{W}$ = 0,1682, $\underline{X}2$ (9) = 18,05 (Siegel, 1956), n'étant pas significatif. En outre, le coefficient de corrélation des rangs moyens de Spearman entre toutes les paires possibles d'utilisateurs n'est que de 0,08, ce qui indique une uniformité négligeable des préférences de classement chez les utilisateurs.

Une analyse de variance avec deux critères de classification de Friedman (Kirk, 1968) indique qu'il n'y a pas de différences significatives entre les préférences des utilisateurs pour ce qui est des pages 1 ($\underline{X}2$ (11) = 18,50, $\underline{p}$ =0,07), aucune page 1 n'ayant été jugée, de façon significative, préférable à une autre. Des analyses distinctes de variance de Friedman concernant les classements établis par les cinq anglophones et les cinq personnes bilingues n'étaient pas plus significatives.

Les analyses, tant de Friedman que de Kendall, indiquent qu'il n'y avait pratiquement pas d'homogénéité entre les initiés quant à la qualité des diverses versions de la page 1 qui leur étaient présentées. On en conclut qu'il n'y a pas de différences marquées entre les différentes versions présentées ou que la notion de la qualité d'une page était passablement différente d'un initié à l'autre. L'expérience suivante fournit des données qui permettent de choisir entre ces deux interprétations.

EXPÉRIENCE 2

La première expérience donnait à entendre que pratiquement tous les initiés étaient d'opinion différente sur les éléments constituant une page d'index ramifié de qualité. Malgré cette divergence, il est plausible que les non-initiés aient, comme groupe, une perception assez homogène de l'arbre. Cependant, il existe deux autres interprétations des résultats de la première expérience: chacun perçoit l'arbre à sa façon ou les douze pages sur lesquelles portent les expériences sont à ce point semblables qu'on ne peut les différencier véritablement. (Cependant, un simple coup d'oeil aux douze pages d'index de la première expérience permet presque d'infirmer cette deuxième interprétation.) La deuxième expérience a permis de retenir une de ces trois interprétations.

La deuxième expérience avait pour objet de déterminer les préférences et le rendement de non-initiés du Télidon relativement aux douze versions de la page 1 d'index utilisées dans la première expérience. Au cours de l'expérience, les sujets avaient à accomplir deux tâches: tout d'abord, chacun d'eux devait faire une recherche documentaire au moyen de chacune des douze versions de la page d'index, puis classer ces pages selon ses préférences. L'expérience visait aussi à déterminer la relation entre le rendement et les préférences des utilisateurs relativement aux index ramifiés.

MÉTHODE

SUJETS. Nous avons demandé à vingt fonctionnaires qui ne connaissaient pas le Télidon de participer à cette expérience. Dix étaient bilingues (parlant le français et l'anglais) et dix ne parlaient que l'anglais. Ces personnes ont été choisies parmi des volontaires.

STIMULI. Pour la deuxième tâche de l'expérience (établissement des préférences), chaque carte de stimulus présentait une version de la page 1. A l'exception de quelques modifications minimes, les stimuli étaient ceux de la première expérience. Voici en quoi consistaient les modifications: a) toutes les rubriques étaient en majuscules et les descripteurs en minuscules; b) l'expression "du guide d'intérêt général" fut ajoutée à la première rubrique de la page réalisée par l'indexateur professionnel; et c) un espace de six caractères fut laissé entre la rubrique et les descripteurs.

Pour la première tâche de l'expérience (recherche et extraction), la forme et la formulation des pages visualisées correspondaient exactement à celles des cartes de stimulus utilisées pour la première tâche de la première expérience.

TÂCHES. Tous les sujets avaient deux tâches à accomplir. Ils devaient d'abord effectuer une recherche documentaire pour trouver des renseignements précis en choisissant la bonne rubrique sur chaque page d'index. Ils devaient ensuite classer les douze versions de la page 1 selon leurs préférences.

La base de données à l'étude au cours de la présente expérience est celle qui a été utilisée par Lee et Latrémouille (1979). Pour obtenir un échantillon représentatif des informations contenues dans

la base de données du Télidon, dix des 900 documents que renferme la base de données ont été choisis au hasard. L'information contenue dans chaque document a servi à formuler une question. Chaque question a été formulée avec soin pour éviter de mettre par inadvertance les utilisateurs sur la bonne voie (voir la liste des questions à l'Annexe B). L'étude n'exigeait pas une recherche dans toutes les ramifications de l'arbre. Seule la première page de chaque index était présentée aux participants. Pour chaque question, ils ne pouvaient choisir qu'une seule rubrique par page. Chaque participant devait répondre à dix questions par page, c'est-à-dire à 120 au total.

Nous n'avons pas indiqué aux participants si leurs choix étaient bons ou mauvais. Les résultats obtenus sont donc un reflet très fidèle des réponses que donneraient des non-initiés auxquels aucun élément de réponse ne serait fourni. Une rétroaction aurait permis aux sujets de tirer profit de leurs erreurs. L'utilisateur qui aurait su qu'une information était classée sous une autre rubrique aurait pu compenser les déficiences de la conception de l'arbre et éviter le même type d'erreur dans les questions suivantes. Les expérimentateurs n'ayant pas fourni d'éléments de réponse aux sujets, le modèle expérimental rend compte de façon optimale des déficiences de conception des pages d'index. Le test portant sur la consultation des pages aurait été beaucoup moins valable si une rétroaction avait permis aux sujets d'éviter de choisir une mauvaise version même si celle-ci leur semblait être la bonne.

DÉMARCHE. Tous les sujets ont d'abord effectué la recherche, puis l'établissement des préférences. Au début de la séance, il y a eu une brève description du Télidon à l'intention des participants. Les douze pages d'index ont ensuite été visualisées une à une dans un ordre aléatoire différent pour chaque sujet. On a demandé à chacun de répondre à la question affichée au haut de l'écran en choisissant la

rubrique appropriée dans la page d'index affichée au-dessous. Pour chaque page un, dix questions étaient posées à chaque utilisateur, dans un ordre aléatoire différent pour chaque sujet. La liste de ces questions est reproduite à l'Annexe B. Les utilisateurs répondaient à chaque question en introduisant le numéro correspondant à la rubrique choisie dans la page d'index. Après quelques instants, une nouvelle question apparaissait à l'écran. Les expérimentateurs n'indiquaient pas aux sujets si leurs choix étaient bons ou mauvais.

Une fois la recherche terminée, on a demandé aux utilisateurs de classer par ordre de préférence les douze versions de la page d'index. La méthode utilisée pour ce faire correspondait essentiellement à celle qui avait été utilisée dans la première expérience.

RÉSULTATS

Même si tous les participants connaissaient assez bien l'anglais, la langue maternelle de la moitié d'entre eux était le français et celle de l'autre moitié l'anglais. Une analyse de variance factorielle à blocs réduits (Kirk, 1968) portant sur des mesures répétées du facteur pages d'index (12 niveaux) et des groupes indépendants pour ce qui est du facteur langue maternelle (2 niveaux: francophones et anglophones) montrent que ni la langue maternelle (F (1, 18) = 2,62, NS), ni l'interaction de la langue maternelle par page d'index (F (11, 198) = 1,76, NS) n'a d'effet sur le rendement. En outre, le coefficient de corrélation des rangs de Spearman entre les préférences moyennes de classement des douze pages par les anglophones et les francophones est de 0,92, p inférieur à 0,01. Étant donné que la variance concernant les préférences de classement des pages par les francophones et par les anglophones est la même dans une proportion de 85%, les sujets des deux catégories ont probablement suivi essentiellement le même raisonnement pour classer les douze pages. Comme la langue maternelle n'a pas eu d'effet appréciable sur les réponses des participants, les données concernant l'ensemble des 20 sujets ont été regroupées aux fins de toutes les analyses ultérieures.

PRÉFÉRENCES

Le tableau 2 donne le rang moyen et le rang normalisé (Garner & Creelman, 1970) de chaque page d'index. (Voir l'Annexe C pour connaître les préférences de classement indiquées par chacun.) Les utilisateurs ont indiqué des préférences passablement homogènes comme le montre le coefficient de concordance de Kendall, W = 0,5195, $X^2(19)$ = 114,29, p inférieur à 0,001 (Siegel, 1956). Le coefficient de corrélation des rangs moyens de Spearman entre les classements par ordre de préférence des différents utilisateurs est de

0,49, ce qui représente un degré relativement élevé d'homogénéité et d'uniformité des non-initiés relativement à leurs préférences pour les diverses pages d'index.

Une analyse de variance avec deux critères de classification de Friedman (Kirk, 1968) indique qu'il y a des différences significatives entre les préférences des utilisateurs pour ce qui est des pages d'index, ($\underline{X}2$ (11) = 132,95, $\underline{p}$ inférieur à 0,001). Le tableau 3 présente les résultats d'un test de comparaison multiple non paramétrique fondé sur les sommes des rangs de Friedman (Hollander & Wolfe, 1973). Cette analyse montre qu'il n'y a pas de différence significative entre les six pages jugées les meilleures et qu'il en va de même des cinq pages jugées les moins bonnes. Cependant, on a jugé que la page classée au premier rang (page modifiée, avec descripteurs) était significativement supérieure aux six pages classées au bas de l'échelle des préférences (voir le tableau 3 pour plus de détails).

Six des douze versions de la page d'index étaient systématiquement différentes l'une de l'autre. Il existait deux versions de la page originale, de la page modifiée et de la page réalisée par l'indexateur professionnel. Une version de chacune des pages contenait des descripteurs après chaque rubrique, tandis que l'autre n'en contenait pas. Le test des signes (Siegel, 1956) est un test statistique de l'impact qu'a sur les préférences des utilisateurs l'adjonction de descripteurs aux pages d'index. Ce test révèle qu'il y a beaucoup plus d'utilisateurs qui préfèrent consulter des pages d'index contenant des descripteurs. Pour ce qui est de la page d'index originale, tous les sujets (20) préféraient la version avec descripteurs, $\underline{p}$ étant bien inférieur à 0,001. D'autre part, dans le cas de la page modifiée et de la page réalisée par l'indexateur professionnel, 16 des 20 utilisateurs préféraient les versions avec descripteurs ($\underline{p}$ = 0,006).

RENDEMENT

Le tableau 4 indique le nombre moyen d'erreurs commises dans chaque page d'index. Il y a un degré élevé d'uniformité quant au nombre d'erreurs commises à chaque page, comme en témoigne le coefficient de concordance de Kendall W = 0,5205, $X^2(19)$ = 114,50, p inférieur à 0,001. Pour toutes les paires possibles de sujets, le coefficient moyen de corrélation des rangs de Spearman entre les classements du nombre d'erreurs par page est relativement élevé comme c'était le cas des préférences des utilisateurs: r de Spearman = 0,50.

Au total, les utilisateurs ont commis en moyenne 46,85 erreurs, ce qui représente une moyenne de 3,90 erreurs par page par utilisateur. Chaque utilisateur pouvait commettre au plus dix erreurs, puisque chacun faisait dix choix par page. Donc, 39% des choix faits par les utilisateurs étaient erronés, ce qui donne une probabilité d'erreur de 0,39 lorsqu'un sujet choisit une rubrique dans une page d'index.

Comme on peut le constater en consultant le tableau 4, des différences significatives existent, pour les douze pages d'index, entre le nombre d'erreurs par page: $F(11, 209)$ = 22.48, p inférieur à 0,001. Le test F modéré de Geisser-Greenhouse (Geisser & Greenhouse, 1958; Greenhouse & Geisser, 1959), pour des degrés modifiés de liberté de 1 et de 19, est aussi très significatif: p inférieur à 0,001. Le test de comparaison multiple de Neuman-Keuls révèle que le nombre moyen d'erreurs est comparable pour les sept pages d'index qui ont fait l'objet du plus petit nombre d'erreurs et que ces sept pages ont été jugées significativement supérieures aux cinq autres pages d'index. Les deux pages contenant le plus d'erreurs par page sont significativement pires que les dix autres (voir tableau 5).

Il y avait, intégré au modèle global de douze pages, un modèle d'analyse de variance portant sur des mesures répétées 2 x 3 des deux facteurs répétitifs suivants: le type de page d'index et les pages avec descripteurs par rapport à celles sans descripteur. L'interaction des descriptions par page est significative puisque $F(2, 95) = 7,10$, p inférieur à 0,01 et que les résultats du test F modéré de Geisser-Greenhouse donnent $F(1, 19) = 7,10$, p inférieur à 0,05. Même si l'interaction est significative, elle rend compte de seulement 4,2% de la variation totale. De plus, Il y a aussi interaction entre les différents genres de pages pour ce qui est du nombre moyen d'erreurs par page (voir figure 2).

Etant donné que l'interaction était significative, nous n'avons fait qu'une analyse simple des effets principaux sur les données (voir tableau 7). Le tableau montre que l'adjonction de descripteurs à chacune des trois pages d'index permet d'obtenir une page de beaucoup améliorée (moins d'erreurs). Le degré d'amélioration varie considérablement d'une page à l'autre mais il est bien supérieur dans le cas de la page d'index originale (voir Annexe A). L'adjonction de descripteurs aux pages d'index a permis des réductions d'erreurs supérieure à 50% dans le cas de la page d'index originale et de près de 50% dans le cas de la page modifiée. Il n'y avait pas de relation systématique entre le degré d'amélioration et la qualité (mesurée par le nombre d'erreurs) d'une page d'index sans descripteurs. Il est donc impossible de prédire dans quelle proportion l'adjonction de descripteurs à une page peut améliorer celle-ci. Cependant, les résultats de la présente expérience semblent confirmer que l'adjonction de descripteurs à une page d'index peut réduire sensiblement le nombre d'erreurs commises par l'utilisateur.

Le test des signes a donné des résultats comparables à l'analyse de variance du modèle 2 x 3. Les utilisateurs ont commis significativement moins d'erreurs lorsque les pages contenaient des

descripteurs. Pour ce qui est de la page originale aussi bien que de la page modifiée, 18 des 20 utilisateurs ont commis moins d'erreurs dans les pages où il y avait des descripteurs que dans les autres, p étant bien inférieur à 0,001 (une égalité). Pour ce qui est de la page réalisée par l'indexateur professionnel, 13 des 20 utilisateurs ont commis moins d'erreurs lorsqu'il y avait des descripteurs, $p = 0,048$ (deux égalités).

Il existe une corrélation étroite entre le classement moyen des douze pages d'index en fonction du nombre d'erreurs par page et le classement moyen des préférences, comme en témoigne la formule de Spearman $r = 0,72$, p inférieur à 0,01. Pour chaque sujet, la corrélation de Spearman a été calculée entre les préférences de classement de l'utilisateur et le nombre d'erreurs que celui-ci a commises à chaque page; la corrélation moyenne des 20 utilisateurs est de 0,38, ce qui est très significatif, p étant inférieur à 0,001. Les cotes moyennes de préférences de classement des initiés lors de la première expérience correspondent significativement à celles des non-initiés (r de Spearman = 0,53, p inférieur à 0,05), mais sont à l'opposé des rendements des non-initiés si l'on tient compte du nombre d'erreurs (r de Spearman = 0,22, NS).

Lors de la séance qui a suivi l'expérience, on a demandé à chaque sujet de dire selon quels critères il avait classé les pages. Le tableau 8 donne la liste de ces critères et leur fréquence. Au total, pour les deux expériences, un ou plusieurs participants ont fait état de sept critères différents, que nous reproduisons ci-après.

1. Ampleur des descriptions - Les participants ont fait des observations sur les descriptions ajoutées en regard des rubriques de certaines pages. Certains ont signalé que la présence de descripteurs ralentissait la recherche étant donné que le

temps de lecture était plus long. D'autres ont indiqué que ces ajouts se traduisaient par un nombre moins élevé d'erreurs au moment de la recherche et qu'ils réduisaient le besoin de revenir au début de l'arbre.

2. Généralité des titres - Les participants jugeaient de la valeur d'une page d'après la généralité des titres. Un titre était considéré comme trop général lorsqu'il ne faisait pas intervenir une notion ou une idée. Par exemple, "DIVERS" entrait dans cette catégorie puisque ce titre peut regrouper à peu près n'importe quoi.

3. Complexité de la terminologie - Les participants ont fait des observations sur le choix des mots utilisés pour désigner les rubriques. Parfois, un titre, quoique juste, peut comporter des mots trop techniques ou trop complexes. "BASES DE DONNÉES D'EXPÉRIENCE-PILOTE" est un titre où les termes sont complexes.

4. Ordre des titres - Certains participants accordaient de l'importance à l'ordre dans lequel se présentaient les rubriques d'une page. Deux personnes qui ont participé à la deuxième expérience ont indiqué que l'ordre de présentation des rubriques d'une page facilitait le cheminement logique de la recherche. Un sujet a indiqué que les premières rubriques de la page devraient être plus générales et les dernières, plus précises, ajoutant qu'il était plus facile d'éliminer d'abord les catégories plus générales. D'autres ont signalé que les rubriques les plus susceptibles d'être consultées et celles qu'il fallait retrouver rapidement ("Urgence", par exemple) devraient figurer en haut de page.

5. Nombre de titres (rubriques) - Certains participants semblaient accorder de l'importance au nombre de titres par page. Ces personnes soutenaient que les recherches étaient plus efficaces lorsqu'une page

contenait de nombreuses rubriques parce que l'utilisateur du système devait alors consulter un nombre plus restreint de pages pour trouver l'information voulue.

6. Répartition de l'information - Les participants ont déclaré qu'ils préféraient les pages présentant de multiples catégories à celles où les renseignements recherchés sont regroupés sous un seul titre. Par exemple, dans le cas d'une page, 90% des réponses aux questions se trouvaient dans la rubrique "Guide d'intérêt général".

7. Contenu des catégories - Certains participants jugeaient de la valeur d'une page d'après le contenu des catégories. Ainsi, certains estiment que l'expression d'une seule idée par catégorie est préférable à la réunion de plusieurs sous un même titre. Par exemple le titre "Affaires et gouvernement" devrait être scindé en deux catégories. Certains n'ont pas évalué la page "Expert 6" de façon favorable. Ils estimaient qu'elle prêtait à confusion, l'accent étant mis sur l'action plutôt que sur le contenu. Ainsi, ce n'est pas nécessairement pour se payer une sortie que l'on consulte la liste des restaurants. Quelqu'un pourrait vouloir dresser la liste des restaurants dans un secteur de la ville en vue de lancer sa propre affaire, par exemple. Le tableau 8 donne la liste des critères d'après lesquels les utilisateurs ont classé les pages et la fréquence de ces critères.

DISCUSSION

La première expérience montrait qu'il n'y avait pour ainsi dire pas de préférence commune entre les initiés pour ce qui est des pages d'index du niveau supérieur qui leur étaient présentées. En revanche, les résultats de la deuxième expérience montrent une homogénéité appréciable entre les non-initiés du Télidon sur le plan des préférences et du rendement relativement aux pages d'index. Pareille similarité signifie que des utilisateurs de milieux différents ont tendance à employer les mêmes critères lorsqu'ils jugent de la qualité des pages d'index. L'étroite corrélation entre le rendement et les préférences donne à entendre que les préférences subjectives se fondent, au moins en partie, sur les difficultés que connaissent les utilisateurs au cours de leur recherche. Donc, si l'on améliore un index ramifié en réduisant le nombre d'erreurs, cette amélioration devrait aussi se traduire par une satisfaction accrue de l'utilisateur lorsqu'il consultera cet index.

La différence entre l'homogénéité d'un groupe (non-initiés) et le manque de cohérence de l'autre (initiés) concernant le classement des douze pages d'index d'après les préférences des utilisateurs peut être attribuable à deux facteurs distincts. Premièrement, par suite de leur expérience du Télidon, les initiés peuvent en être venus à percevoir de façon très différente les structures ramifiées. Deuxièmement, ces initiés auraient peut-être eu des préférences beaucoup plus homogènes s'ils avaient pu, comme l'autre catégorie de participants, chercher un renseignement donné dans chaque page. Les données recueillies par la présente étude ne permettent pas de savoir quelle est, de ces deux interprétations, celle qu'il faut retenir.

Dans l'ensemble, la meilleure page 1 est la page modifiée qui contient des descripteurs. C'est celle que les utilisateurs ont préféré le plus et qui est venue au deuxième rang pour ce qui est

du nombre d'erreurs. (La page d'index dans laquelle le moins d'erreurs ont été commises -page expert 4 avec un seul descripteur- n'est venue qu'au sixième rang quant aux préférences des utilisateurs, rang qui est significativement inférieur à celui de la page modifiée avec descripteurs.) Deux facteurs peuvent expliquer la supériorité de cette dernière. En premier lieu, la page modifiée représente, par rapport à la page d'index originale conçue par Lee & Latrémouille (1981), une amélioration fondée sur des procédés empiriques. Dans le cas de la série d'expériences menées par Lee & Latrémouille (1981), la structure ramifiée originale a été testée dans des conditions plus réalistes, les erreurs ont été localisées et elles ont été corrigées. On pouvait donc s'attendre à ce que la page d'index modifiée constitue une amélioration. En second lieu, l'adjonction de descripteurs à la page modifiée en a amélioré la conception de façon significative. Cinq des douze pages d'index figurent parmi les pages jugées les meilleures quant aux deux mesures et les différences entre les deux ne sont pas significatives, ni sur le plan des préférences, ni sur celui du rendement: modifiée:D, originale:D, expert 2:D, expert 3:1D et expert 1:ND. Quatre de ces cinq pages contiennent des descripteurs, ce qui confirme davantage la supériorité des pages contenant des descriptions par rapport à celles qui n'en contiennent pas. De plus, compte tenu des résultats de la présente expérience, on pourrait très probablement améliorer de façon sensible la cinquième page (expert 1:ND) en y ajoutant des descripteurs.

Une constatation importante de la présente étude, c'est la valeur manifeste de l'adjonction de descripteurs à une page d'index. Non seulement l'adjonction de descripteurs à une page se traduit-elle par une amélioration significative dans chaque cas, mais le degré d'amélioration est lui aussi important. Il y a eu deux ou trois fois plus d'erreurs commises dans les pages sans descripteurs. De même, sur le plan des préférences, les utilisateurs ont classé ces pages à un rang bien inférieur comparativement à celui des pages avec

descripteurs. Une autre indication de l'importance de descripteurs sur la préférence des utilisateurs est présentée à la figure 3. Dans une étude précédente, (Lee & Latrémouille, 1981), la page originale sans descripteurs a été testée (comme l'ont été les autres pages de la structure ramifiée originale), les erreurs ont été repérées puis corrigées pour produire l'index ramifié "modifié", version améliorée. La figure 3 montre que la modification de la page originale sans descripteur (aux fins de produire une page modifiée sans descripteur) s'est traduite par une amélioration significative sur le plan du classement moyen. Cependant, l'amélioration du classement moyen de la page originale à laquelle on a simplement ajouté des descripteurs a été bien supérieure à l'amélioration qui a résulté de la correction des erreurs décelées dans l'étude précédente de Lee & Latrémouille (1981).

Les commentaires des non-initiés quant à leurs critères de classement des douze pages d'index sont utiles pour les concepteurs d'index ramifiés du vidéotex puisqu'elles leur indiquent les facteurs qui jouent dans la satisfaction des utilisateurs concernant la structure arborescente. Le nombre de descripteurs est le critère qui revient le plus souvent. Presque tous ceux qui ont indiqué ce critère (65% des utilisateurs) préféraient des rubriques avec descripteurs. Cependant, quelques-uns ont ajouté que ces descripteurs devaient être courts et concis. Un seul utilisateur a dit préférer des rubriques sans descripteur.

La généralité des rubriques est le deuxième critère le plus fréquent (60% des non-initiés). Des rubriques comme "guide d'intérêt général" et "divers" ont été jugées trop ambiguës et de nature trop générale. Pour le non-initié, ces titres sont sans valeur, car ils ne reflètent nullement le type d'informations auxquelles ils renvoient. Comme autre critère, les non-initiés ont indiqué, dans une proportion de 45%, l'ordre ou la suite des rubriques sur une page d'index.

Cependant, on a constaté certaines différences quant aux différents ordres que préféraient les utilisateurs. Plusieurs estiment qu'il faut placer les rubriques les plus consultées au haut de la page de façon à réduire le plus possible le temps de consultation. D'autres croient qu'il faut placer les rubriques les plus générales au haut de la liste et les plus précises au bas de la liste. Finalement, 35% des utilisateurs ont indiqué que le nombre de rubriques par page doit être assez élevé, considérant que trois ou quatre rubriques par page constituent une perte de temps puisqu'il faut alors consulter un plus grand nombre de pages. Ceci confirme donc la recommandation de Lee (1979) selon laquelle il faut éviter de placer seulement deux ou trois rubriques par page puisque les temps de recherche sont alors trop longs.

Compte tenu des observations faites par les non-initiés, nous pouvons faire les recommandations qui suivent à l'intention des concepteurs de structures arborescentes du vidéotex. Tout d'abord, il est avantageux d'avoir recours à des descripteurs. Deuxièmement, il faut éviter l'utilisation de titres trop généraux et ambigus comme "guide d'intérêt général" et "divers". Troisièmement, le nombre de rubriques par page ne doit pas être trop restreint (2, 3 ou 4). Quatrièmement, si c'est possible, il faut utiliser des rubriques à peu près équivalentes quant à la quantité d'informations auxquelles chacune renvoie (plus particulièrement, on doit éviter les titres comme "guide d'intérêt général" qui contiennent de 80% à 90% de toutes les données qu'il est possible d'extraire d'une page d'index donnée). Cinquièmement, les indexateurs doivent, dans la mesure du possible, disposer les rubriques en commençant par la plus générale et en terminant par la plus précise, ou les classer selon la fréquence de consultation, en commençant par celle qui est consultée le plus souvent.

A partir des résultats de la présente étude, nous recommandons aux responsables de bases de données du vidéotex d'étudier la possibilité d'ajouter des descripteurs aux pages d'index, particulièrement à celles des niveaux supérieurs de l'arbre. Cette mesure accroîtrait sans doute sensiblement la satisfaction de l'utilisateur tout en réduisant de beaucoup le nombre d'erreurs. Si des descriptions sont utiles aux non-initiés parce qu'elles clarifient le type d'informations répertoriées sous les rubriques, elles devraient aussi être utiles à tous les niveaux de la structure arborescente. Cependant, d'autres recherches seront nécessaires pour savoir s'il est nécessaire d'adjoindre des descripteurs aux rubriques de tous les niveaux de la base de données. Il reste aussi à déterminer l'effet de descripteurs sur les temps de transmission, d'affichage, de lecture et de recherche.

OUVRAGES DE RÉFÉRENCE

Garner, W.R., & Creelman, C.D., Problems and methods of psychological scaling. Tiré de G.F. Sumers (Ed.) Attitude measurement. Chicago, Rand McNally, 1970.

Hollander, M., & Wolfe, D.A., Nonparametric statistical methods. New York: Wiley, 1973, chapitre 7.

Kirk, R., Experimental design: Procedures for the behavioural sciences. Belmont, Californie: Brooks/Cole, 1968.

Latrémouille, S., & Lee, E., The design of Telidon tree indexes: Improving tree indexes by testing naive users. Présenté pour publication sous forme de note technique, ministère des Communications, Canada, 1981.

Lee, E., & Latrémouille, S., Évaluation de la structure arborescente de l'information dans la base de données du Télidon. (Note technique no BRG 79-12). Ministère des Communications, Ottawa. Décembre 1979. Aussi disponible dans Recherche en comportement sur Télidon 1. Ministère des Communications, Ottawa, 1980, pp. 228 à 236.

McEwen, S., Étude de l'efficacité de la recherche documentaire au moyen du système Télidon. Présenté pour publication sous forme de note technique, ministère des Communications, Canada, 1981.

Siegel, S., Nonparametric statistics for the behavioural sciences. New York: McGraw-Hill, 1956.

Van Ness, F.L., & Tromp, J.H., Is view data easy to use? Institute for Perception Research. Eindhoven, Pays-Bas: IPO Annual Progress Report, nº 14, 1979.

Whalen, T., & Latrémouille, S., L'efficacité d'un index ramifié lorsque l'existence de données est incertaine. (Note technique nº BRIC 80-3). Ministère des Communications, Ottawa. Septembre 1980.

Tableau 1

PRÉFÉRENCES DE CLASSEMENT DES PAGES D'INDEX
PAR LES INITIÉS: EXPÉRIENCE 1

RANG	PAGE D'INDEX[a]	RANG NORMALISÉ[b]	RANG MOYEN[b]
1	Modifiée:ND	14,82 (0,69)	2,90 (2,08)
2	Modifiée:D	9,40 (1,07)	5,10 (3,87)
3	Expert 2:ND	6,25 (1,28)	5,80 (4,49)
4	Expert 1:ND	4,47 (0,76)	6,40 (2,76)
5	Expert 3:1D	4,37 (0,92)	6,40 (3,50)
6	Originale:ND	4,03 (0,78)	6,20 (2,97)
7	Expert 5:ND	3,78 (0,77)	6,95 (3,32)
8	Expert 4:1D	3,40 (0,74)	6,90 (3,11)
9	Réalisée par un indexateur professionnel:ND	2,60 (0,70)	7,25 (2,66)
10	Originale:D	1,47 (0,93)	7,50 (3,57)
11	Expert 6:D	1,37 (1,24)	8,05 (4.26)
12	Réalisée par un indexateur professionnel:D	0 (0,54)	8,55 (2,29)

[a] D = page d'index avec descripteurs; ND = page d'index sans descripteurs; 1D = page d'index dont une seule rubrique contient un descripteur.

[b] Les inscriptions entre parenthèses représentent l'écart-type des résultats pour la moyenne correspondante.

Tableau 2

PRÉFÉRENCES DE CLASSEMENT DES PAGES UN D'INDEX PAR LES NON-INITIÉS: EXPÉRIENCE 2

RANG	PAGE D'INDEX[a]	RANG NORMALISÉ[b]	RANG MOYEN[b]
1	Modifiée:D	43,65 (0,40)	2,40 (1,05)
2	Expert 2:D	42,38 (0,92)	3,05 (3,14)
3	Originale:D	40,28 (0,52)	3,00 (1,72)
4	Expert 1:ND	27,34 (0,48)	5,40 (2,01)
5	Modifiée:ND	26,93 (0,86)	5,85 (3,01)
6	Expert 3:1D	24,53 (0,82)	5,60 (2,85)
7	Expert 4:1D	22,81 (0,64)	6,25 (2,49)
8	Réalisée par un indexateur professionnel:D	15,26 (0,46)	7,95 (1,73)
9	Expert 6:D	10,48 (0,67)	8,65 (2,21)
10	Originale:ND	8,25 (0,40)	9,40 (1,50)
11	Réalisée par un indexateur professionnel:ND	3,76 (0,56)	10,05 (1,96)
12	Expert 5:ND	0 (0,65)	10,40 (2,06)

[a] D = page d'index avec descripteurs; ND = page d'index sans descripteurs; 1D = page d'index dont une seule rubrique contient un descripteur.

[b] Les inscriptions entre parenthèses représentent l'écart-type des rangs pour la moyenne correspondante.

Tableau 3

TESTS DE COMPARAISON MULTIPLE PORTANT SUR LES DIFFÉRENCES ENTRE LES SOMMES DES RANGS JUGÉS PRÉFÉRABLES POUR LES PAGES D'INDEX: EXPÉRIENCE 2[a]

Page d'index[b]	Page d'index[b] 10	5	4	6	2	7	12	9	1	3	8
11	12	13	60	64	69	77*	*	*	*	*	*
10	-	1	48	52	57	65	99*	*	*	*	*
5		-	47	51	56	64	98*	*	*	*	*
4			-	4	9	17	51	65	80*	*	*
6				-	5	13	47	61	76*	*	*
2					-	8	42	56	71	84*	*
7						-	34	48	63	76*	*
12							-	14	29	42	49
9								-	15	28	35
1									-	13	20
3										-	7
8											-

[a] Les inscriptions du tableau représentent les différences entre les sommes des rangs des pages d'index indiquées dans les rangées et les colonnes correspondantes.

[b] Page 1 = page originale:ND; 2 = page modifiée:ND; 3 = page réalisée par un indexateur professionnel:ND; 4 = expert 1:ND; 5 = expert 2:D; 6 = expert 3:1D; 7 = expert 4:1D; 8 = expert 5:ND; 9 = expert 6:D; 10 = page originale:D; 11 = page modifiée:D; 12 = page réalisée par un indexateur professionnel:D

* Les inscriptions comportant un astérisque sont significativement différentes pour p inférieur à 0,05.

Tableau 4

NOMBRE MOYEN D'ERREURS COMMISES PAR LES NON-INITIÉS A CHAQUE PAGE UN DE L'INDEX: EXPÉRIENCE 2

RANG	PAGE D'INDEX[a]	NOMBRE MOYEN D'ERREURS[b]
1	Expert 4:1D	2,25 (2,00)
2	Modifiée:D	2,50 (1,76)
3	Originale:D	2,65 (2,43)
4	Expert 2:D	2,75 (1,52)
5	Expert 3:1D	2,75 (1,89)
6	Expert 1:ND	3,05 (2,16)
7	Expert 6:D	3,25 (1,59)
8	Expert 5:ND	4,60 (1,96)
9	Modifiée:ND	4,85 (1,39)
10	Réalisée par un indexateur professionnel:D	5,35 (2,30)
11	Originale:ND	6,35 (2,28)
12	Réalisée par un indexateur professionnel:D	6,50 (1,76)

[a] D = page d'index avec descripteurs; ND = page d'index sans descripteur; 1D = page d'index dont une seule rubrique contient un descripteur.

[b] Les inscriptions entre parenthèses représentent l'écart-type des rangs pour les moyennes correspondantes.

Tableau 5

ANALYSE DE COMPARAISON MULTIPLE DE NEUMAN-KEULS CONCERNANT LE NOMBRE MOYEN D'ERREURS PAR PAGE D'INDEX: EXPÉRIENCE 2[a,b]

		Page d'index										
		11	10	5	6	4	9	8	2	12	1	3
Page d'index	7	0,25	0,40	0,50	0,50	0,80	1,00	2,35*	2,60*	3,10*	4,10*	4,25*
	11	-	0,15	0,25	0,25	0,55	0,75	2,10*	2,35*	2,85*	3,85*	4,00*
	10		-	0,10	0,10	0,40	0,60	1,95*	2,20*	2,70*	3,70*	3,85*
	5			-	0	0,30	0,50	1,85*	2,10*	2,60*	3,60*	3,75*
	6				-	0,30	0,50	1,85*	2,10*	2,60*	3,60*	3,75*
	4					-	0,20	1,55*	1,80*	2,30*	3,30*	3,45*
	9						-	1,35*	1,60*	2,10*	3,10*	3,25*
	8							-	0,25	0,75	1,75*	1,90*
	2								-	0,50	1,50*	1,65*
	12									-	1,00*	1,15*
	1										-	0,15
	3											-

[a] Page 1 = page originale:ND; 2 = page modifiée:ND; 3 = page réalisée par un indexateur professionnel:ND; 4 = expert 1:ND; 5 = expert 2:D; 6 = expert 3:1D; 7 = expert 4:1D; 8 = expert 5:ND; 9 = expert 6:D; 10 = page originale:D; 11 = page modifiée:D; 12 = page réalisée par un indexateur professionnel:D

[b] Les inscriptions du tableau représentent les différences entre les moyennes des pages d'index indiquées dans les rangées et les colonnes correspondantes.

* Les inscriptions comportant un astérisque sont significativement différentes pour p inférieur à 0,05.

Tableau 6

ANALYSE DE MESURES RÉPÉTÉES DU MODÈLE INTÉGRÉ 2 X 3
(PAGES X DESCRIPTEURS): Expérience 2

SOURCE	df[a] (degrés de liberté)	SS (somme des carrés)	MS (carré moyen)	F	VARIATION TOTALE EN %
Sujets soumis aux mêmes conditions	19	248,53			
Pages	2(1)	103,65	51,82	22,62**	13,4%
Descripteurs	1(1)	172,80	172,80	75,42**	22,3%
Pages X descripteurs	2(1)	32,55	16,28	7,10*	4,2%
Erreur	95(19)	217,67	2,29		
TOTAL	119	775,20			

[a] Les degrés de liberté du test F modéré de Geisser-Greenhouse sont indiqués entre parenthèses.

*p inférieur à 0,01 pour le test classique et p inférieur à 0,05 pour le test F modéré.

**p inférieur à 0,001 pour les tests F classique et modéré.

Tableau 7

ANALYSE SIMPLE DES EFFETS PRINCIPAUX DU MODÈLE INTÉGRÉ 2 X 3 (DESCRIPTEURS X PAGES): EXPÉRIENCE 2

SOURCE	df (degrés de liberté)	SS (somme des carrés)	MS (carré moyen)	F
SUJETS SOUMIS AUX MÊMES CONDITIONS	19	248,53		
Pages (P)	2	103,65	51,82	22,62**
Avec descripteurs	1	102,90	102,90	22,47**
Sans descripteur	1	33,30	33,90	7,27**
Descripteurs (D)	1	172,80	172,80	75,42**
Page originale D	1	136,90	136,90	59,78**
Page modifiée D	1	55,22	55,22	24,11**
Page réalisée par un indexateur professionnel D	1	13,22	13,22	5,77*
Descripteurs X pages	2	32,55	16,28	7,10**
Erreur	95	217,67	2,29	
TOTAL	119	775,20		

* p inférieur à 0,05
** p inférieur à 0,01

Tableau 8

FRÉQUENCE DES CRITÈRES DU CLASSEMENT PAR ORDRE DE PRÉFÉRENCE DES 12 PAGES D'INDEX: EXPÉRIENCES 1 et 2

	Expérience 1 (N=10)	Expérience 2 (N=20)
1. Quantité de description		
Favorable:	5	10
Favorable, mais avec titres plus concis:	-	2
Contre:	5	1
2. Généralité des titres	7	12
3. Complexité des termes	2	6
4. Ordre des titres	5	9
5. Nombre de rubriques par page		
Devrait être faible	1	-
Devrait être élevé	3	7
6. Répartition de l'information	2	5
7. Contenu des catégories	6	9

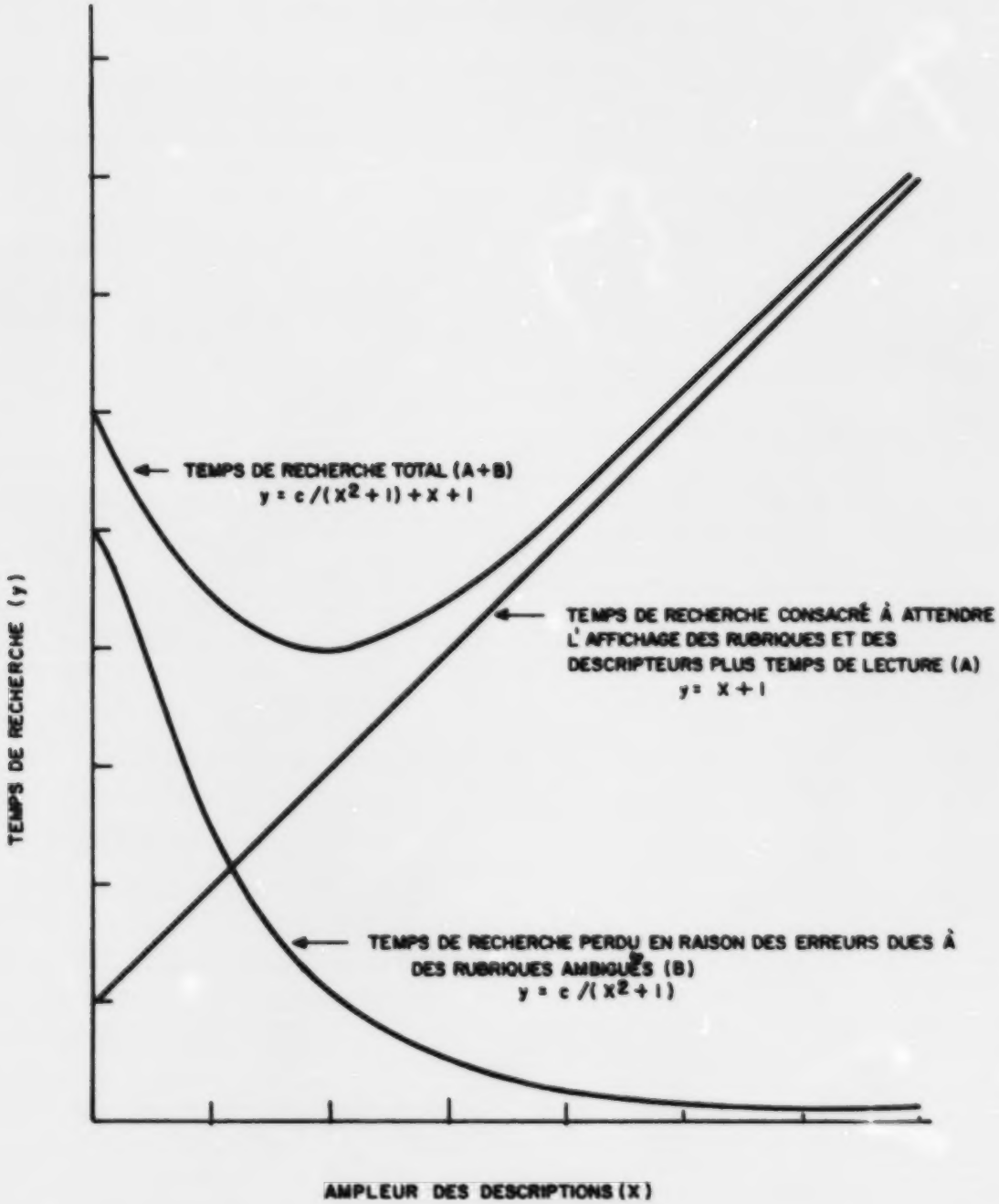

FIG. 1 RELATION PRÉVUE ENTRE L'AMPLEUR DES DESCRIPTIONS ET LE TEMPS DE RECHERCHE DE L'UTILISATEUR

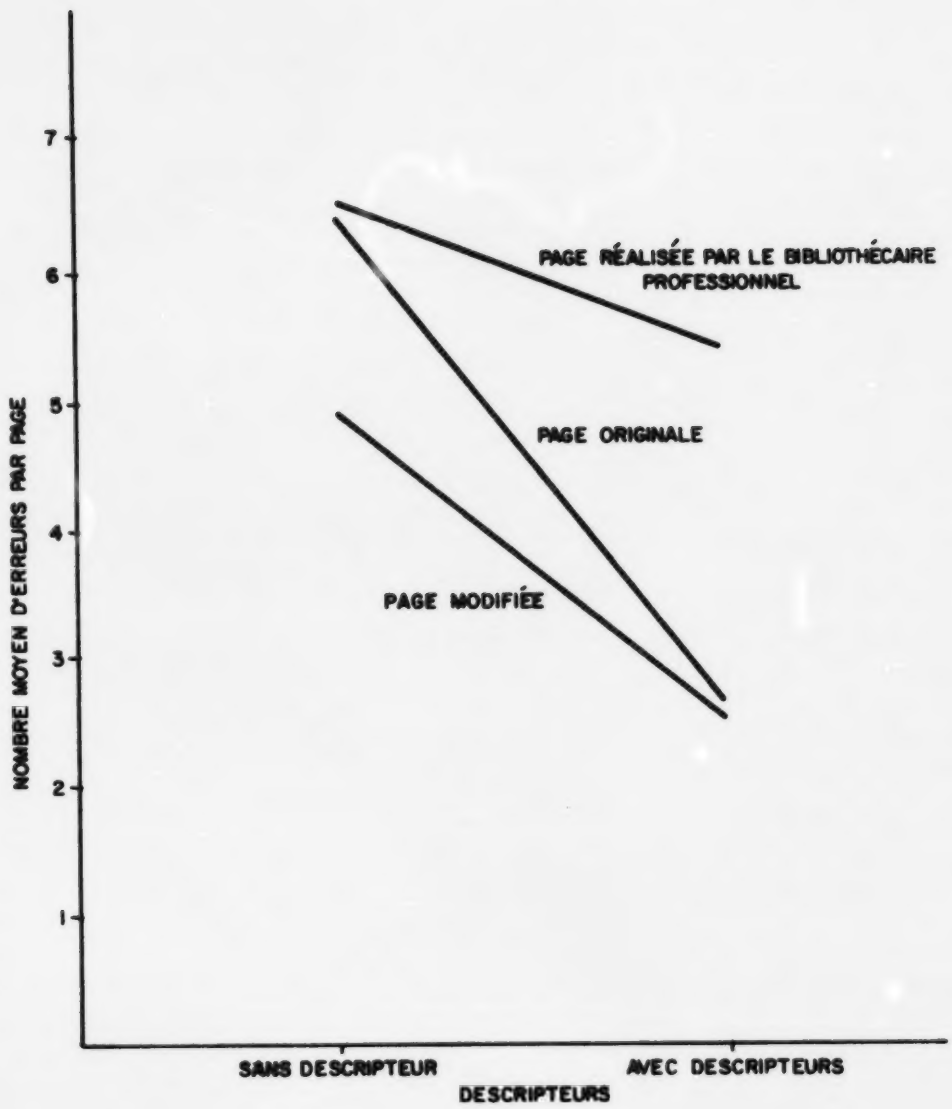

FIG. 2 INTERACTION DESCRIPTEUR-PAGE-EXPÉRIENCE 2

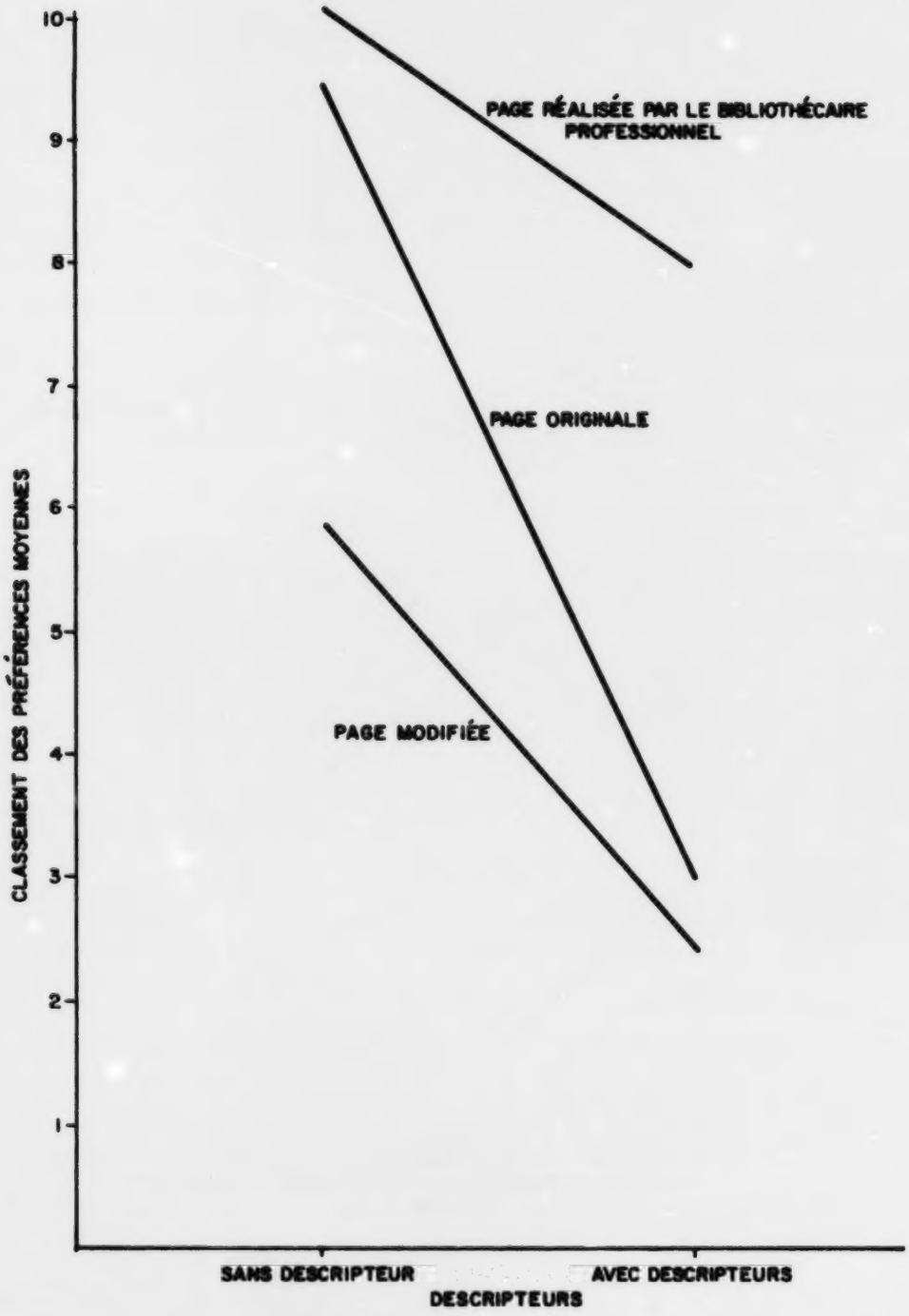

FIG. 3 COMPARAISON DES PRÉFÉRENCES MOYENNES DE CLASSEMENT DES PAGES D'INDEX AVEC DESCRIPTEURS ET DES PAGES D'INDEX SANS DESCRIPTEURS - EXPÉRIENCE 2

ANNEXE A

VOICI LES 12 PAGES UN D'INDEX QUI ONT SERVI AUX FINS DES EXPÉRIENCES 1 et 2.

PAGE ORIGINALE:ND

1. GUIDE D'INTÉRÊT GÉNÉRAL
2. GUIDE DES AFFAIRES
3. GOUVERNEMENT DU CANADA
4. URGENCE
5. GUIDE DE L'UTILISATEUR
6. EXPLICATION DU TÉLIDON
7. NUMÉROS DE TÉLÉPHONE

PAGE MODIFIÉE:ND

1. EXPLICATIONS ET INSTRUCTIONS CONCERNANT LE TÉLIDON
2. URGENCE
3. RENSEIGNEMENTS SUR LES AFFAIRES, LES FINANCES ET LE GOUVERNEMENT
4. DIVERTISSEMENTS ET VOYAGES
5. RENSEIGNEMENTS GÉNÉRAUX
6. ACHATS, VENTES, LOCATIONS ET SERVICES
7. ÉDUCATION
8. EMPLOI
9. ANNUAIRE TÉLÉPHONIQUE

PAGE RÉALISÉE PAR LE BIBLIOTHÉCAIRE: ND

1. GUIDE D'INTÉRÊT GÉNÉRAL
2. AFFAIRES
3. DIVERS
4. URGENCE
5. TÉLIDON
6. TABLEAU D'AFFICHAGE
7. NUMÉROS DE TÉLÉPHONE

EXPERT 1:ND

1. NOUVELLES, MÉTÉO ET SPORTS, AVIS
2. DIVERTISSEMENTS ET RÉCRÉATION
3. CONSOMMATION ET AFFAIRES
4. GOUVERNEMENT DU CANADA
5. SANTÉ, ÉDUCATION, BIEN-ÊTRE, EMPLOI

6. URGENCES
7. COMMENT UTILISER LE TÉLIDON
8. FONCTIONNEMENT DU TÉLIDON (TECHNIQUE)
9. ANNUAIRE TÉLÉPHONIQUE
10. VOYAGES

EXPERT 2:D

1. URGENCE: Numéros de téléphone, premiers soins
2. ACTUALITÉS: Nouvelles, météo, sports, nouvelles locales
3. RENSEIGNEMENTS GÉNÉRAUX: Bulletins sur la consommation, conversion au système métrique, éducation, guide de santé
4. SECTION DES AFFAIRES: Achats, ventes, locations, services commerciaux, questions financières, conseils sur les affaires
5. VOUS AVEZ BESOIN D'UN EMPLOI?
6. VOUS PARTEZ EN VOYAGE?: Transports, voyages, guide touristique
7. DÉTENTE: Sorties, détente chez soi, événements spéciaux, Sports
8. GOUVERNEMENT DU CANADA: Chambre des Communes, députés, rapports du gouvernement fédéral
9. RENSEIGNEMENTS SUR LE TÉLIDON: Descriptions techniques et mode d'utilisation
10. NUMÉROS DE TÉLÉPHONE

EXPERT 3:1D

1. NOUVELLES, MÉTÉO ET SPORTS
2. DIVERTISSEMENTS ET VOYAGES
3. ÉDUCATION ET OUVRAGES DE RÉFÉRENCE
4. CONSOMMATION
   a. Achats et services locaux
   b. Guide du consommateur
   c. Emploi
   d. Immobilier (logement)
5. AFFAIRES ET GOUVERNEMENT
6. URGENCE ET SOINS MÉDICAUX
7. BASES DE DONNÉES D'EXPÉRIENCE-PILOTE
8. EXPLICATION ET INSTRUCTIONS CONCERNANT LE TÉLIDON

EXPERT 4:1D

1. GUIDE D'INTÉRÊT GÉNÉRAL
   Nouvelles, météo et sports
   Divertissements (locaux)
   Achats - Logement
   Éducation
   Annonces quotidiennes
   Voyages et tourisme
   Ouvrages de référence
   Emploi

2. AFFAIRES ET FINANCIAL POST
3. LE GOUVERNEMENT DU CANADA
4. BASES DE DONNÉES D'EXPÉRIENCE-PILOTE
5. URGENCE ET SOINS MÉDICAUX
6. INSTRUCTIONS ET EXPLICATIONS CONCERNANT LE TÉLIDON

EXPERT 5:ND

1. ACTIVITÉS PERSONNELLES
2. AFFAIRES
3. GOUVERNEMENT
4. NOUVELLES, MÉTÉO, SPORTS
5. URGENCE
6. POUR COMPRENDRE LE TÉLIDON

EXPERT 6:D

1. SORTIES: Divertissements locaux, agences de voyage, achats, etc.
2. S'OCCUPER CHEZ SOI: Jeux éducatifs, histoires, rénovations, etc.
3. ACTUALITÉS: Nouvelles, résultats sportifs, prévisions de la météo, etc.
4. OUTILS DE RÉFÉRENCE: Numéros de téléphone, encyclopédies, etc.

PAGE ORIGINALE:D

1. GUIDE D'INTÉRÊT GÉNÉRAL: Nouvelles, météo, sports, divertissements, consommation, emploi, voyages, loisirs, conseils, éducation, tableau d'affichage
2. GUIDE DES AFFAIRES: Les affaires nationales, liste des entreprises, placements, carrières, occasions, marché de l'immobilier, publicité, commercialisation
3. GOUVERNEMENT DU CANADA: Chambre des Communes, sénat, rapports du gouvernement fédéral, gouvernements provinciaux, administrations municipales, premiers ministres
4. URGENCE: Numéros de téléphone, organismes, premiers soins
5. GUIDE DE L'UTILISATEUR: Comment se servir du Télidon
6. EXPLICATION DU TÉLIDON: Explication technique
7. NUMÉROS DE TÉLÉPHONE

PAGE MODIFIÉE:D

1. EXPLICATIONS ET INSTRUCTIONS CONCERNANT LE TÉLIDON
2. URGENCE: Numéros de téléphone et premiers soins
3. RENSEIGNEMENTS SUR LES AFFAIRES, LES FINANCES ET LE GOUVERNEMENT: Questions financières, placements, économie, impôts
4. DIVERTISSEMENTS ET VOYAGES: Divertissements chez soi et à l'extérieur, sports, transports, guide touristique, agences de voyage
5. RENSEIGNEMENTS GÉNÉRAUX: Information du consommateur, conversion au système métrique, horoscope, nouvelles, météo, sports, avis
6. ACHATS, VENTES, LOCATIONS ET SERVICES: Achat de marchandises ou d'une propriété, location d'équipement ou d'appartements, services commerciaux par exemple rénovation de la maison, plomberie, électricité
7. ÉDUCATION
8. EMPLOI
9. ANNUAIRE TÉLÉPHONIQUE

PAGE RÉALISÉE PAR LE BIBLIOTHÉCAIRE:D

1. TABLE DES MATIÈRES du guide d'intérêt général
2. AFFAIRES: Rapports, occasions d'affaires, immeubles commerciaux, services commerciaux
3. DIVERS: Tables de conversion au système métrique, Horoscope...
4. URGENCE: Numéros de téléphone, Premiers soins
5. TÉLIDON: Comment s'en servir? Comment fonctionne-t-il?
6. TABLEAU D'AFFICHAGE: Naissances, Décès, Annonces personnelles
7. NUMÉROS DE TÉLÉPHONE

ANNEXE B

QUESTIONS

1. Trouvez une liste de téléviseurs à vendre.

2. Quels sont les moyens de transport individuel disponibles?

3. Quelles sont les prévisions météorologiques locales pour aujourd'hui?

4. Quelles qualités doit posséder une cuisinière?

5. Où se trouve le restaurant 'CHINESE VILLAGE'?

6. Trouvez le jeu mathématique "Quel est le numéro manquant".

7. Trouvez la façon de porter plainte lorsqu'une entreprise vous a escroqué.

8. Trouvez des renseignements sur le restaurant afghan 'KHYBER PASS'.

9. Quelles qualités doit posséder un congélateur?

10. Découvrez ce que sont les exportations mondiales de céréales.

ANNEXE C

CLASSEMENT DES PAGES EN FONCTION DES PRÉFÉRENCES INDIVIDUELLES

EXPÉRIENCE 1

		Page											
		1	2	3	4	5	6	7	8	9	10	11	12
Sujet	#1	2	8	5	6	9	10	4	7	12	11	3	1
	2	8	4	2	3	12	10	7	1	9	11	6	5
	3	8	6	3	1	9,5	9,5	7	2	4	5	11,5	11,5
	4	12	3	6	1	8	7	9	2	4	5	11	10
	5	5	11	1	10	6	3	2	12	4	9	7	8
	6	4	6	1	9	3	11	5	10	8	7	2	12
	7	9	3	6	1	8	10	5	2	7	4	11	12
	8	4	12	3	9	6	8	7	2	11	10	5	1
	9	5	11	1	9	4	8	6	12	2	3	7	10
	10	5	11	1	2	7	9	12	8	3	4	6	10

Note. Pages: 1 = page originale:ND, 2 = page originale:D, 3 = page modifiée:ND, 4 = page modifiée:D, 5 = page réalisée par un indexateur professionnel:ND, 6 = page réalisée par un indexateur professionnel:D, 7 = expert 1:ND, 8 = expert 2:D, 9 = expert 3:1D, 10 = expert 4:1D, 11 = expert 5:ND, 12 = expert 6:D

EXPÉRIENCE 2

Sujet	Page 1	2	3	4	5	6	7	8	9	10	11	12
#1	10	1	8	2	12	9	7	3	5	4	11	6
2	8	3	4	1	10	7	5	2	12	6	11	9
3	10	1	6	5	11	9	7	3	4	2	12	8
4	10	2	8	3	11	7	9	1	4	5	12	6
5	8	3	5	2	11	7	6	1	4	12	10	9
6	11	3	1	2	12	7	4	6	5	10	9	8
7	9	5	6	4	12	8	2	1	3	7	11	10
8	11	1	6	3	9	12	4	2	5	7	8	10
9	8	2	11	3	10	6	7	1	5	4	9	12
10	11	3	8	1	10	9	7	2	5	4	12	6
11	10	3	8	1	11	6	9	2	4	5	12	7
12	10	2	4	1	11	9	5	6	3	7	8	12
13	11	2	9	3	10	4	6	1	5	8	12	7
14	11	2	8	3	10	7	6	1	5	4	12	9
15	10	8	9	2	11	7	3	1	5	4	12	6
16	7	5	1	3	8	9	2	11	12	10	4	6
17	7	5	1	2	9	10	3	4	6	8	11	12
18	11	2	7	3	10	9	6	1	4	5	12	8
19	8	3	6	2	10	9	5	1	4	7	11	12
20	7	4	1	2	3	8	5	11	12	6	9	10

Note. Pages: 1 = page originale:ND, 2 = page originale:D, 3 = page modifiée:ND, 4 = page modifiée:D, 5 = page réalisée par un indexateur professionnel:ND, 6 = page réalisée par un indexateur professionnel:D, 7 = expert 1:ND, 8 = expert 2:D, 9 = expert 3:1D, 10 = expert 4:1D, 11 = expert 5:ND, 12 = expert 6:D